ANALYSIS OF THE PHYSIOLOGICAL ROLES OF THE NONSENSE MEDIATED MRNA DECAY PATHWAY IN *DROSOPHILA MELANOGASTER*

by

Alex Chapin

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Human Genetics

The University of Utah

August 2014

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Alex Chapin**

has been approved by the following supervisory committee members:

Mark Metzstein	, Chair	**6/6/2014** Date Approved
Carl Thummel	, Member	**6/6/2014** Date Approved
Mark Yandell	, Member	**6/6/2014** Date Approved
Michael Howard	, Member	**6/6/2014** Date Approved
Julie Hollien	, Member	Date Approved

and by **Lynn Jorde** , Chair/Dean of

the Department/College/School of **Human Genetics**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

The nonsense mediated mRNA decay (NMD) pathway is a conserved posttranscriptional mRNA decay pathway that functions to destabilize a variety of naturally occurring target mRNAs. The NMD pathway functions in all eukaryotes and regulates a significant portion of the transcriptome. It is thought that this regulation is critical as inhibition of NMD leads to physiological and developmental defects in all organisms and in the case of more complex organisms, lethality. It is predicted that overexpression of NMD pathway target genes leads to these defects in NMD mutants. Despite the critical nature of this pathway, little is know about how NMD functions in a developmental and physiological context, including which target genes are most critically regulated by NMD and how the overexpression of these targets may mediate the NMD mutant phenotype. To address this knowledge gap, we first use two genome-wide techniques to identify and characterize the kinds of transcripts targeted by NMD in the context of an intact metazoan, *Drosophila melanogaster.* We then examine more closely the function of one of these target genes, *Gadd45*, and find that overexpression of this target in NMD mutants may explain important aspects of the NMD mutant phenotype.

This thesis is dedicated to my parents, S. Kim Chapin

and Anne R. Constable

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Two well-studied aspects of regulated gene expression are the mechanisms of transcriptional activation and transcriptional repression (Shandilya and Roberts, 2012). While transcriptional control does indeed contribute to the regulation of many genes, the steady-state level of any given mRNA is a combination of not only its transcription rate, but also the rate of RNA degradation. In many cases, gene expression is primarily regulated by modulation of mRNA decay. This process is executed by a number of pathways, some of which function in the decay of all transcripts, while others, such as the nonsense mediated mRNA decay (NMD) pathway, target a select set of transcripts (Garneau et al., 2007; Kervestin and Jacobson, 2012). Through its ability to regulate the expression of these specific transcripts, NMD has important cellular and organismal functions (Hwang and Maquat, 2011).

NMD is a conserved cellular process that functions in all eukaryotes examined. This pathway is named for its ability to target and degrade mRNA alleles harboring nonsense mutations while passing over cognate wild-type messages (Chang et al., 2007). However, NMD can also degrade other RNAs, including certain transcripts that contain no mutation whatsoever (Dahlseid et al., 1998). The *trans*-acting factors that execute

NMD are generally required for normal development and physiology, suggesting that the negative regulation of native RNAs targeted by the NMD pathway is a critical feature of eukaryotic biology (Behm-Ansmant et al., 2007). However, the full extent of NMD-mediated regulation, in particular which target genes are most critically regulated by NMD, remain unknown.

To analyze the role of NMD in development and physiology *in vivo*, we have exploited genetic techniques available for the fruit fly, *Drosophila melanogaster.* This has allowed us to address many outstanding questions such as 1) Are there specific NMD-target genes that mediate the NMD loss-of-function phenotype?, 2) What portion of the transcriptome is regulated by NMD?, 3) What *cis*-acting features of native target mRNAs allow for recognition by the NMD machinery?, and 4) Does NMD have any tissue-specific roles?

Introduction to NMD

The term "nonsense mediated decay" was coined in 1993 by Allan Jacobson to describe the phenomena in which decay rates of nonsense-mutation bearing transcripts are increased over those of wild-type transcripts (Belasco, 1993). In general, this accelerated decay leads to lower steady-state levels of mutant transcripts and thus fewer molecules of encoded polypeptide (Chang et al., 2007). NMD was first described in *S. cerevisiae* (Losson and Lacroute, 1979) but has since been shown to operate in all eukaryotes examined, including humans (Maquat et al., 1981). In certain cases, NMD function has been demonstrated by direct measure of the differential RNA stability between wild-type and nonsense-mutant transcripts (Losson and Lacroute, 1979; Maquat

et al., 1981). In other cases, NMD has been demonstrated to function indirectly, through the steady-state accumulation of target transcripts when NMD activity is inhibited.

NMD is executed by a suite of evolutionarily conserved *trans*-acting proteins that function together in an "NMD pathway" (Behm-Ansmant et al., 2007). The best-studied NMD pathway genes are the *Upfs* (for Up-frameshift suppressor, (Culbertson et al., 1980)) and the *Smgs* (for Suppressor with morphogenic effects on genitalia, (Hodgkin et al., 1989)). *Upf1, Upf2,* and *Upf3* were discovered as required for nonsense-transcript turnover in yeast (Leeds et al., 1991, 1992), and homologues with equivalent roles were subsequently identified in all eukaryotes. Suppressor screens in *C. elegans* identified the homologues of all three *Upf* genes called *smg-2*, *smg-3*, and *smg-4*, for *Upf1, Upf2,* and *Upf3*, respectively (Hodgkin et al., 1989). In addition, four other genes, *smg-1, smg-5, smg-6,* and *smg-7* were identified (Cali et al., 1999; Hodgkin et al., 1989). These latter *Smg* genes are found throughout metazoa, although plants may have a distant homologue of *Smg7* and *Smg1* . More recently discovered NMD pathway genes include *Smg8* and *Smg9*, identified as binding partners of SMG1 (Yamashita et al., 2009), and *Smgl1* and *Smgl2*, which are required to degrade an NMD-sensitive transgenic reporter gene (Longman et al., 2007). These factors are poorly conserved, and their roles in NMD are less well understood.

Mechanisms of NMD targeting

NMD substrate recognition is thought to function through the ability of the NMD machinery to recognize inherent differences between translation termination that occurs at nonsense, or premature, termination codons (PTCs) caused by nonsense mutations, and

termination at natural termination codons (NTCs) (Ghosh et al., 2010; Kervestin and Jacobson, 2012; Kervestin et al., 2012). Consistent with this notion, NMD is translation dependent and the application of pharmacological inhibitors that inhibit translation elongation or initiation block NMD and stabilize PTC-bearing transcripts (Carter et al., 1995; Zhang et al., 1998). Additionally, ribosomal toeprinting analysis, a technique that maps sites of ribosomal pausing during translation, reveals unreleased pausing at PTCs in nonsense mutation containing mRNAs that are targeted by NMD (Amrani et al., 2004). Pausing at PTCs is thought to be indicative of inefficient release of ribosomal subunits following termination (Amrani et al., 2004), a process that is so rapid at NTCs that no pausing can be detected (Cui et al., 1995; Pisarev et al., 2010).

Following the aberrant termination that occurs at PTCs, NMD targeting is initiated through binding of the central *trans*-acting effector, UPF1, an ATP-dependent RNA helicase whose helicase function is required for NMD (Chamieh et al., 2008). UPF1 connects translation termination to NMD by interacting with both the terminating ribosome and the target RNA (Ivanov et al., 2008). The helicase domain of UPF1 directly interacts with the target mRNA (Bhattacharya et al., 2000), while other regions of the protein bind the ribosome indirectly via interactions with the eukaryotic release factors, eRF1 and eRF3 (Ivanov et al., 2008). Formation of the UPF1/release factor complex is stimulated by UPF1 binding to UPF2 (Mendell et al., 2000), as well as by interactions between UPF2 and UPF3 (Chamieh et al., 2008; Maderazo et al., 2000). In metazoans, UPF1 activity is further regulated by the SMG proteins, with UPF1 being phosphorylated by the phosphatidylinositol 3-kinase-related kinase (PIKK) family protein, SMG1 (Yamashita et al., 2001). Along with UPF1 and the release factors, SMG1 forms a stable

complex, termed the SURF complex (for SMG1, UPF1 and release factors) (Kashima et al., 2006). Formation of the SURF complex promotes UPF1 phosphorylation and is thought to be a prerequisite for NMD function (Yamashita et al., 2009). Phosphorylated UPF1 can then bind SMG6 (Okada-Katsuhata et al., 2012), an endoculease, which has been shown to cleave target mRNAs near the PTC (Eberle et al., 2009; Huntzinger et al., 2008). To initiate dephosphorylation of UPF1, phosphorylated UPF1 recruits a complex of the proteins SMG5 and SMG7, which in turn recruit the phosphatase PP2A (Anders et al., 2003). Dephosphorylation by PP2A is thought to facilitate dissociation of NMD proteins and the targeting of new mRNAs, allowing for recycling of NMD proteins onto new target mRNAs.

In all eukaryotes, decay of NMD targets is driven by recruitment of the decapping complex, consisting of DCP1 and DCP2, to the 5' end of the transcript (Lejeune et al., 2003; Mitchell and Tollervey, 2003). The decapping complex removes the 5' 7-methylguanosine cap of the mRNA, exposing the 5' phosphate to the ubiquitous cytoplasmic 5' to 3' exonuclease, XRN1 (Arribas-Layton et al., 2013). Additionally, in metazoans, decay is initiated by endonucleolytic cleavage of target mRNAs by SMG6 (Gatfield and Izaurralde, 2004; Huntzinger et al., 2008), which exposes an XRN1-sensitive 5' phosphate and a 3' hydroxyl, which is susceptible to the cytoplasmic exonuclease, the exosome (Houseley et al., 2006).

Recognition of substrate RNAs

One of the most intriguing aspects of NMD function is how the NMD machinery is able to distinguish target from nontarget mRNAs. This problem has mostly been

studied by examining the difference between wild-type and nonsense-bearing NMD-targeted transcripts. These studies have lead to two nonexclusive models that explain how the NMD pathway might recognize nonsense transcripts. One of these models, the "*faux* 3' UTR" model, is based primarily on work in yeast. Work in mammalian cell culture has lead to the "exon junction complex (EJC)" model (Amrani et al., 2004)(Figure 1.1).

The faux 3' UTR model is based on the observation that the 3' UTRs of nonsense transcripts are inherently different and longer than those of the cognate wild-type 3' UTR. This model states that under normal circumstances, translation termination and ribosomal reloading onto the 5' end of the transcript are promoted by specific interactions that occur between the natural 3' UTR/poly A tail, or associated proteins, and the terminating ribosome. These interactions do not take place when translation terminates at a PTC, which leads to inefficient ribosome dissociation and the ability of the NMD machinery to initiate decay. Proteins associated with mRNAs, known as ribonuclear proteins (RNPs), such as poly A binding protein (PABP), associate with the poly A tail. PABP has been shown to promote efficient translation termination via interactions with the release factors, interactions that are mutually exclusive with UPF1 binding (Amrani et al., 2004; Kervestin et al., 2012). Thus, the faux 3' UTR model states that RNPs (such as PABP) compete with UPF1 for binding release factor, and when the interaction with UPF1 dominates, NMD is initiated. Consistent with this, when the long(er) 3' UTRs of nonsense transcripts are experimentally shortened, bringing the poly A tail into closer proximity with the termination codon, transcripts are rendered insensitive to NMD (Eberle et al., 2008). Similarly, long 3' UTRs that are NMD-sensitive can be rendered insensitive by the experimental introduction of double-stranded foldback motifs that also

bring the poly A tail into closer proximity with the termination codon (Eberle et al., 2008).

From work in mammalian cell culture, a different model has been put fourth that emphasizes the importance of a specific kind of RNP. In all metazoans, an assembly of proteins, known as the exon junction complex (EJC), associates with mRNA within 25 nt upstream of exon-exon junctions concurrent with splicing in the nucleus (Le Hir et al., 2000). During the first, or "pioneer" round of translation, EJCs are removed or remodeled (Gehring et al., 2009; Sato and Maquat, 2009). Under the EJC model, if a ribosome encounters a termination codon with an EJC positioned downstream, NMD is triggered (Nagy and Maquat, 1998). This model is supported by evidence in mammals that NMD can only occur during the pioneer round of translation (Hwang et al., 2010; Maquat et al., 2010). Furthermore, for many mammalian genes, NMD does not target transcripts in which the PTC is in a terminal exon. This is also found for PTCs that are positioned in the penultimate exon, but are still downstream of the site of EJC deposition. Also supporting this model is the observation that a majority of wild-type transcripts harbor NTCs in the terminal exon and are not targeted (Nagy and Maquat, 1998). Moreover, the NMD factor UPF3 is a component of the EJC, and UPF2 also associates with the EJC once the transcript is exported to the cytoplasm (Chang et al., 2007).

While the exact proposed mechanisms appear different, the faux 3' UTR and EJC models essentially state that local RNP interactions around the site of translation termination are used by the NMD machinery to detect termination at PTCs. In addition, NMD in mammals is not exclusively driven by detection of downstream EJCs. NMD can also be initiated by long 3' UTRs even in the absence of splice junctions (and thus no

downstream EJCs). These observations imply that downstream EJCs are not absolutely necessary to initiate NMD in mammals and therefore, the EJC may represent a mechanism that potentiates NMD, but is not the underlying trigger mechanism.

Another sequence feature thought to help specify NMD targets is the so-called downstream sequence element (DSE), located in 3' UTRs. The DSE was identified in yeast as a 12-base pair *cis*-acting sequence that, when deleted, partially stabilizes certain NMD target transcripts (Peltz et al., 1993). DSEs may stimulate NMD by recruiting the protein HRP1, which has been shown to bind both DSEs and UPF1 *in vitro* (González et al., 2000). In this way, HRP1 acts analogously to EJCs to stimulate NMD when translational termination occurs at a PTC (González et al., 2000). DSE-like sequences have not been reported in metazoans, although it is possible that other similar *cis*-acting features exist.

NMD pathway function is essential for normal development and physiology

The conservation of NMD pathway genes throughout eukaryota suggests that these genes perform essential functions. The requirement for NMD genes could be explained by two general models. First, NMD genes could be required because function of the NMD pathway itself is necessary for development and viability (Hwang and Maquat, 2011; Palacios, 2013). Conversely, genes involved in the NMD pathway could have requirements in non-NMD functions. In support of the first model is the observation that loss of individual NMD genes leads to remarkably similar phenotypes in the same organism (Table 1.1). In *S. cerevisiae*, disruption of any *Upf* gene results in cold

sensitivity (Altamura et al., 1992). In *S. pombe*, loss of the Upf genes leads to sensitivity to oxidative stress (Rodríguez-Gabriel et al., 2006). In *C. elegans,* mutations in any of the seven *Smg* genes gives rise to abnormal genital development (Cali et al., 1999; Hodgkin et al., 1989). In *Drosophila*, null mutants of *Upf1* and *Upf2* have an identical lethal phase, with the majority of larvae dying in the second larval instar (Chapin et al., 2014). In zebrafish, morpholino-mediated depletion of *Upf1, Upf2, Smg6,* or *Smg5* lead to greater than 80% reduction in viability at 5 days post fertilization as well as impaired development of the brain and eye and abnormal somites (Wittkopp et al., 2009). In mouse, loss of *Upf1* (also called *Rent1*), *Upf2,* or *Smg1* results in inviability by embryonic day 7.5, 9.5, and 12.5, respectively (McIlwain et al., 2010; Medghalchi et al., 2001; Weischenfeldt et al., 2012). In *Arabidopsis*, *Upf1* and *Smg7* are both required during early embryogenesis (Riehs et al., 2008; Yoine et al., 2006). Moreover, in mutants in which NMD is only partially inactive, the strengths of other phenotypes are also reduced. For instance, *Drosophila Smg1* mutants have both a mild NMD defect and are mildly affected in terms of overall viability, while null alleles of *Smg6* impart an intermediate defect in both NMD and viability (Frizzell et al., 2012). Taken together, these data suggest it is the function of NMD itself that is required for normal development and physiology in eukaryotes.

Through analyses in many different organisms, the NMD pathway has been implicated as functioning in numerous biological processes. Perhaps the most well studied is regulation of the cell cycle. In *Drosophila*, siRNA-mediated knockdown of any of the six NMD factors in S2 cell culture results in arrest at the G2/M transition of the cell cycle (Rehwinkel et al., 2005). Mosaic analysis reveals that loss of *Upf1* or *Upf2* in

follicle cells results in a cell-autonomous proliferation defect that can be recued by expression of the apoptosis inhibitor, *p35*, suggesting that NMD-deficient cells activate pro-apoptotic pathways (Avery et al., 2011). Mosaic analysis in the eye suggests the proliferation defect of NMD-deficient cells is caused by an inability to effectively compete with wild-type cells in the same field. Cell/cell competition is defined as the cell-cell interactions that sense metabolic state relative to neighboring cells (Johnston, 2009). In a competitive context, cells that have a weakened metabolic state are eliminated through apoptosis. These states are communicated via signaling cues that either promote apoptosis in "losing" cells or compensatory growth in "winning" cells. A defect in cell/cell competition, as opposed to a defect in overall cell viability, can be revealed using a special type of mosaic analysis, the *GMR:hid* technique (Stowers and Schwarz, 1999). Using this technique, NMD mutant cells can be generated while simultaneously eliminating competing wild-type cells through forced induction of apoptosis. In this context, NMD mutant clonal patches expand to sizes much larger than those generated using standard clone-making techniques in which wild-type cells are still present (Metzstein and Krasnow, 2006). These results suggest that a functioning NMD pathway is required to promote a metabolic state that allows cells to effectively compete, rather than NMD being absolutely required for proliferation or viability.

In mammalian cell culture models, NMD genes are required for efficient progression through S-phase, suggesting a role in DNA synthesis (Azzalin and Lingner, 2006). In mice, conditional disruption of *Rent2* (*Upf2*) in the hematopoietic lineage results in a loss of proliferative cells, but has no effect on differentiated, postmitotic cells. Furthermore, murine blastocysts mutant for *Rent1* (*Upf1*) or *Smg1* demonstrate

widespread activation of apoptosis, as measured by TUNEL staining (McIlwain et al., 2010; Medghalchi et al., 2001). Part of the dependence on NMD for proper cell cycle control could be related to a defect in telomere maintenance, as has been demonstrated in human cells silenced for NMD components (Azzalin et al., 2007). In *S. cerevisiae*, NMD components have also been shown to have an epistatic relationship with the telomere maintenance genes *Yku70* and *cdc13*, suggesting that this aspect of NMD function is conserved (Addinall et al., 2008, 2011).

NMD has also been shown to have specific functions in neuronal tissue. From observations in mammal cell culture (Giorgi et al., 2007) and *Drosophila* (Barbee et al., 2006), NMD components are found associated with dendritic mRNPs; complexes of mRNA and protein found in dendrites. In neurons, dendritic mRNPs function in the transport of specific mRNAs to synapses, where the cargo mRNA undergoes local translation. When NMD is inhibited in rat PC-12 cells, via depletion of the EJC factor eIF4III, there is an increase in the amplitude of mini excitatory postsynaptic currents (mEPSC), which is indicative of an increase in the numbers of excitatory receptors found in postsynaptic regions of dendrites (Giorgi et al., 2007). This modulation of receptor levels may be related to the function of the immediate early gene, *arc*, which is a natural target of NMD. *Arc* mRNA is a dendritic mRNP cargo message, and ARC protein has been shown to modulate receptor levels as part of the process of long-term potentiation at synapses (Plath et al., 2006). NMD has also been shown to function at the synapse in *Drosophila*, as mutations in *Smg1* (also known as *nonC*), *Upf2,* or *Smg6* all lead to abnormal morphology at the neuromuscular junction (NMJ) and impaired synaptic vesicle cycling (Long et al., 2010).

NMD has also been shown to have a functional role in brain development in mammals. In the brains of mice, *Upf1* mRNA is negatively regulated by expression of a microRNA, *miR-128,* that binds in the 3' UTR of *Upf1*. *miR-128* expression increases over the course of murine embryonic development such that NMD activity in the brain is gradually repressed (Bruno et al., 2011). This increase in *miR-128* activity over development is in agreement with data from cell culture-based studies, which suggest expression of *miR-128* (and thus regulated suppression of NMD) is required for proper neuronal differentiation (Bruno et al., 2011). NMD has also been shown in mammals to function in the regulation of axon guidance. NMD components are found in the growth cones of various classes of murine neurons, and disruption of NMD results in disruption of axon trajectories during development (Colak et al., 2013). The role of NMD in this process may be related to an mRNP cargo message, *robo3.2*, a natural NMD target (by virtue of a retained intron that creates a PTC in this isoform) that is locally translated in growth cones (Colak et al., 2013). Evidence from human studies also suggests a role for NMD in the central nervous system. Loss of function mutations in one of the two homologues of *Upf3*, *Upf3B* (or *Upf3X*), are associated with two forms of X-linked mental retardation, Lujan-Fryns syndrome and FG syndrome (Tarpey et al., 2007). Both these syndromes are characterized by intellectual deficits, implying that NMD could also play a functional role in human neurodevelopment.

Two models to explain eukaryotic dependence on NMD

The precise mechanisms of the dependence on NMD for normal development and physiology in eukaryotes are not well understood. However, NMD has been shown in

many cases to target and degrade various naturally occurring RNAs ("native" RNAs), in addition to its classically studied targets, nonsense-mutation containing transcripts. It is likely that regulation of native RNAs is essential for the various biological processes that rely on NMD. These native target genes fall into two general categories. The first type is "junk" RNA, i.e, RNAs that do not code for any obviously functional protein nor carry out any other productive function in the cell. For instance, NMD has been shown to degrade unprocessed mRNAs in the cytoplasm, intragenic transcripts, and transcripts of viral origin (Neu-Yilik et al., 2004). In contrast, NMD has also been shown to directly target fully processed and otherwise fully functional mRNAs. Thus, it appears that NMD can act as a cellular surveillance pathway that degrades unproductive junk RNAs and as a specialized posttranscriptional regulatory pathway with specific target genes. Based on these ideas, two general models have been developed to describe the *in vivo* function of NMD: The "Vacuum Cleaner" model represents the surveillance capabilities of NMD, whereas the "Swiss Army Knife" model emphasize the role of NMD in targeted posttranscriptional gene regulation (Neu-Yilik et al., 2004). The vacuum cleaner model states that NMD is essential due to its surveillance capabilities. Therefore in NMD mutants, organisms succumb to a buildup of unproductive transcripts. The Swiss Army Knife model states that NMD is vital due to its role in negative regulation of specific, wild-type transcripts, and in NMD mutants it is the stabilization of these specific mRNAs that interferes with normal development and physiology.

Chapters 3 and 4 of this dissertation describe experiments that functionally test both the vacuum cleaner and Swiss Army Knife models of NMD function in *Drosophila*. A prediction of the latter model is that NMD mutants succumb to overexpression of a

small number of direct NMD targets. To test this prediction, we conducted a rescue screen described in Chapter 3 that was designed to identify genes whose overexpression suppresses an NMD mutant phenotype. In addition, the Swiss Army Knife model predicts that NMD may have tissue-specific functions due to the targeting of specific transcripts with spatially restricted expression patterns. According to the vacuum cleaner model, NMD would be required in all tissues at all times in development, as presumably all tissues produce aberrant transcripts. To test this distinction, we performed a tissue-specific rescue screen, described in Chapter 4, to look for tissues where NMD function is sufficient for organismal viability.

Identification of NMD pathway target mRNAs

The identification of native target genes is critical to understanding why organisms rely on a functioning NMD pathway. As described by the vacuum cleaner model, NMD targets many types of aberrant RNAs. For instance, incompletely spliced transcripts that escape into the cytoplasm could contain PTCs, thus triggering NMD when such transcripts undergo translation. Indeed, such transcripts were among the first native RNAs to be identified as targets (He et al., 1993). In addition, transcripts emanating from nonfunctional programmed rearrangement (such as in the TCR locus in mammalians) (Wang et al., 2002), unproductively spliced ribosomal proteins (Mitrovich and Anderson, 2000), noncoding rRNAs (Marquardt et al., 2011), intragenic transcripts originating from read-through (Thompson and Parker, 2007), transcripts of viral or transposon origin (Guan et al., 2006), genes that undergo aberrant 3’ UTR length extension (Muhlrad and Parker, 1999), and genes that undergo leaky scanning (Welch and Jacobson, 1999) have

all been identified as NMD targets.

Another large class of unproductive native transcripts that appear to be under regulation by NMD are those created through the process of alternative splicing. Alternative splicing is a broadly employed mechanism to control gene expression and increase transcript diversity that relies on differential splicing to control the relative amounts of various RNA isoforms transcribed from a single locus (Ben-Dov et al., 2008). In humans and *Drosophila*, 85% and 50% of genes are estimated to undergo alternative splicing, respectively. In certain cases, alternative splicing is utilized to regulate the overall abundance of a single productive mRNA isoform, rather than to produce multiple productive forms. By controlling the relative levels of the productive versus unproductive forms, overall levels of the productive form can be regulated. When gene expression is controlled in this way, it is referred to as regulated unproductive splicing and translation (RUST) (Lareau et al., 2007a), and NMD is predicted to degrade unproductive isoforms that generate PTCs (Lewis et al., 2003; McGlincy and Smith, 2008). For example, 50% of transcripts from the mammalian FGFR2 gene contain a PTC but are degraded to undetectable levels in the presence of a functioning NMD pathway (Jones et al., 2001). A similar situation also occurs in the case of the *Drosophila* sex determination gene *tra*, which undergoes an alternative splicing event in females such that the major isoform contains several PTCs creating a NMD-sensitive substrate (Longman et al., 2007; Metzstein and Krasnow, 2006).

Other types of native targets exist that contain special sequence features or utilize special types of regulation that lead to targeting by the NMD pathway. For example, genes that use +1 frameshifting to regulate protein expression (Gatfield and Izaurralde,

2004), transcripts containing uORFs (Welch and Jacobson, 1999), genes with unusually long 3' UTRs (Kebaara and Atkin, 2009), and in mammals, genes with the NTC not positioned in the terminal exon (Mendell et al., 2004) have all been shown to sensitize transcripts to NMD. For example, genes like *Oda* utilize a +1 frameshift as part of normal regulation (Ivanov et al., 1998). In this case, full-length protein is only produced when a ribosomal frameshift to the +1 reading frame is induced by high levels of cytoplasmic polyamines. When this shift does not occur, the result is early termination, which creates PTCs and renders the transcript NMD-sensitive (Gatfield and Izaurralde, 2004). Another feature that has been shown to sensitize native transcripts to NMD is a long natural 3' UTR (Kebaara and Atkin, 2009). Such 3' UTRs are downstream from NTCs, but are longer, on average, than typically found in 3' UTRs. Targeting of these transcripts by the NMD pathway is thought to work in a manner similar to that of nonsense targeting, in that the long natural 3' UTR is recognized as aberrant, and therefore termination at the NTC initiates decay.

Genome-wide approaches to studying NMD targeting

While various attributes of NMD targets have been shown to bestow NMD-sensitivity, none are completely predictive of whether any particular mRNA is targeted, with some transcripts containing what appear to be NMD-sensitizing features, yet these transcripts are not destabilized by NMD. This inconsistency suggests our current knowledge of NMD target gene recognition is incomplete. Thus, the goal of identifying all native NMD target genes cannot by met with a purely predictive bioinformatic approach. Therefore, assays have been developed that rely on various functional

definitions of NMD targets, and when coupled with technologies that query genome-wide, they can identify native targets across the transcriptome.

The most conceptually straightforward and most commonly used technique to identify NMD targets is the identification of genes that are upregulated in NMD-defective organisms or cells. This approach relies on using transcript upregulation as a proxy readout for stabilization, as direct RNA targets should be stabilized when NMD is inhibited, allowing steady-state levels of these transcripts to increase. This type of analysis is also useful for revealing the full extent of NMD-dependent gene expression since upregulation in an NMD-mutant background does not solely identify direct targets of NMD. For instance, levels of secondary targets, which are positively regulated by direct targets, should also be upregulated when NMD is inhibited. Using upregulation in an NMD-mutant background will also identify genes that are dependent on the function of *trans*-acting NMD factors, but not on the process of NMD itself. However, the inclusion of this type of false positive can be minimized by studying sets of genes regulated by multiple NMD factors. Genetic mutation in (or silencing of) any given NMD component generally results in the upregulation of about 10% of the transcriptome, as has been demonstrated in yeasts (He et al., 2003; Matia-González et al., 2013), plants (Rayson et al., 2012), *C. elegans* (Ramani et al., 2009), *Drosophila* (Metzstein and Krasnow, 2006; Rehwinkel et al., 2005), zebrafish (Wittkopp et al., 2009), mouse (Weischenfeldt et al., 2008), and human cells (Mendell et al., 2004; Yepiskoposyan et al., 2011). In a few studies, sets of genes have been defined that are upregulated in multiple backgrounds lacking NMD. In *Arabidopsis*, *Upf1, Upf3,* and *Smg7* negatively regulate a shared set of 206 transcripts (0.9% of all genes). In *S. cerevisiae* strains carrying

mutations in any of the three *Upf* genes, 10 to 12 % of expressed genes show differential upregulation, which represents a set of 792 genes (11.6% of expressed transcripts) (He et al., 2003). In *Drosophila* S2 cells, six well-studied NMD components, *Upf1, Upf2, Upf3, Smg1, Smg5,* and *Smg6,* regulate a common set of 184 genes, representing 3.4% of detectable transcripts (Rehwinkel et al., 2005). Finally, inhibition of *Upf1* or *Smg6* in HeLa cells results in the upregulation of a common set of 158 genes (Yepiskoposyan et al., 2011).

A second method used to enrich for direct targets of NMD relies on techniques that can resolve expression changes between multiple mRNA splice isoforms from a single locus. The logic of this technique is based on the assumption that any secondary transcriptional effects caused by NMD disruption will alter the expression of all isoforms equally, whereas for direct targets, alternatively spliced forms will be differentially affected. Not only is isoform-specific analysis a useful way to measure direct targeting by NMD, but it can also be used to estimate the overall role of NMD in the decay of transcripts produced by RUST. Estimates from *C. elegans* indicate that 25% of all AS events introduce a PTC and that NMD can target a significant portion (59%) of these PTC-bearing, AS-generated mRNAs (Barberan-Soler et al., 2009). In *Drosophila*, a much smaller portion of AS events were identified as NMD-regulated (43 confirmed events from a possible 2244) (Hansen et al., 2009). This disparity between the worm and fly might be due to stricter statistical cutoffs used in these two analyses or the vastly larger sample size used in the *Drosophila* experiment. An important consideration when interpreting the results of experiments examining AS is that NMD has been shown to regulate the expression of splicing factors (Lareau et al., 2007b; Morrison et al., 1997; Ni

et al., 2007). Therefore, it is possible that NMD could indirectly influence the relative expression levels of splice isoforms, which would result in the inclusion of false positives when assigning whether a specific isoform is directly targeted or not.

Other methods to detect direct NMD targeting use biochemical techniques, such as crosslinking-immunoprecipitation (CLIP) coupled with array-based or deep sequencing technology to identify transcripts bound by NMD factors on a genome-wide scale (Hurt et al., 2013; Johansson et al., 2007; Matia-González et al., 2013). Such studies were based on initial observations from *C. elegans* that suggested that SMG-2 (UPF1) preferentially associates with PTC-harboring transcripts when compared to wild-type transcripts (Johns et al., 2007). By inference, native transcripts that are enriched for bound UPF1 are likely to be directly targeted. This has been shown to be the case in yeast (Johansson et al., 2007), where TAP-tagged UPF1 is found bound to 44% of genes that are upregulated in Upf1 mutants, suggesting these transcripts represent true direct targets. A similar analysis in murine embryonic stem cells (mESCs) found that 35% of genes upregulated following siRNA depletion of Upf1 showed enrichment of Upf1 binding (Hurt et al., 2013). UPF1 binding to substrate mRNAs is not, however, an absolute means of target identification as UPF1 has also been shown bound to nontargeted substrates (Hogg and Goff, 2010; Zünd et al., 2013).

Still other methods to identify direct targets genome-wide rely on the decay kinetics of NMD target transcripts, in that direct NMD targets decay at a faster rate compared to nontargets. In a simple application of this principle, timecourse expression profiling has been performed on wild-type and *Upf1Δ* yeast cells following inhibition of transcription using thiolutin (Guan et al., 2006). This analysis revealed that, of genes

upregulated upon NMD inhibition, 46% (279/607) show slower decay rates. Interestingly, of the 279 genes identified as direct targets using this method, only 108 (39%) are part of the core set of direct NMD targets identified using UPF1-binding and a reactivation assay (described below) to define direct targets (Johansson et al., 2007). Another example of an analysis based on altered target stability in NMD mutants has been performed using the BRIC-seq technique to monitor turnover of mRNAs (Tani et al., 2012b). In HeLa cells silenced for Upf1, 23% of the upregulated genes were shown to also be stabilized compared to controls (Tani et al., 2012a).

A different technique that also takes advantage of the decay kinetics of NMD target genes uses acute activation of the NMD pathway, rather than acute inhibition of transcription, to reveal decay kinetics (Johansson et al., 2007; Maderazo et al., 2003). A "reactivation" assay can be executed by expressing a gene product in a cognate NMD mutant and monitoring expression levels of putative target genes over a short time course. Using this technique, it is possible to distinguish direct from indirect targets. Direct targets of NMD respond quickly to this reactivation and thus decay quickly, whereas levels of indirect targets take longer to respond. This technique has been used in *S. cerevisiae*, in conjunction with UPF1 CLIP analysis, to assess the degree of direct targeting (Johansson et al., 2007). Of the 792 genes upregulated at steady state in *Upf1* mutants, 230 (30%) demonstrated both acute reactivation and enrichment of UPF1 binding. In Chapter 2 of this thesis, we adapt this method for use in whole *Drosophila* larvae, the first time such a method has been deployed in metazoa.

Features and functions of the NMD-regulated transcriptome

Using the techniques described above, several studies have sought to describe NMD targets both in terms of the structures or *cis*-targeting features that make transcripts susceptible to NMD and the functions of those target genes. These studies have used various model organisms and different experimental paradigms, making direct comparisons difficult. Nonetheless, a few general themes have emerged. While single NMD genes are required for proper expression of about 10% of the transcriptome, anywhere from 60% to less than 10% appear to be directly regulated. This proportion of direct targets to total indirect targets generally tracks with the complexity of the organism. In *S. cerevisiae*, the vast majority of expression level changes in the three *Upf* mutants result in up, but not downregulation (85 to 96%). This indicates that secondary target effects on overall gene-expression levels, which are predicted to result in equal levels of up and down regulation, are minimal. Furthermore, pairwise comparisons of the genes upregulated between the three NMD mutants indicate a minimum correlation coefficient of 0.96, indicating that the three *Upf* genes in yeast have a very minimal impact on gene expression outside of their roles in NMD. Additionally, 29% of the 792 genes upregulated in the three *Upf* mutants are identified as targets in two completely different assays, the *Upf1* binding assay and a *Upf2* reactivation assay, described earlier.

In more complex organisms, directly regulated NMD targets represent a smaller portion of the transcriptome found upregulated by inhibition of particular NMD factors. For instance, in HeLa cells, just 23% of genes upregulated upon *Upf1* knockdown are found to be stabilized by the same treatment. As direct targets of NMD are predicted to be stabilized by inhibition of NMD, of which only 23% of upregulated genes are direct

targets, indicates that the majority of gene expression changes that occur following *Upf1* silencing are indirect. This small portion of direct targets could be explained by the greater complexity of mammalian physiology compared to yeasts, and thus a greater influence of NMD on secondary targets. Additionally, when compared to yeast, there is generally a greater portion of genes that are downregulated by NMD disruption in mammals, again suggesting that more secondary targets exist in mammals.

Interestingly, some of the techniques used to detect genome-wide NMD targets have generated counterintuitive results. For instance, NMD substrates have been identified that behave like direct substrates in kinetic assays, but are not upregulated in steady state (Johansson et al., 2007; Matia-González et al., 2013; Tani et al., 2012a). In yeast, 55% of genes that decay upon *Upf2* reactivation are not upregulated at steady state. In HeLa cells, 90% of the transcripts that are stabilized in *Upf1*-silenced cells are not upregulated at steady state. These results could imply that a majority of direct NMD substrates are subject to additional regulation that can renormalize expression levels when NMD activity is compromised. An alternative hypothesis is that NMD, or NMD components, can also function as transcriptional activators. Thus, when NMD is inhibited, these transcriptional functions are also disrupted, along with the defects associated with transcript decay. These effects would cancel each other out, leading to little change in steady state levels. Such a model is supported by evidence that other mRNA decay factors are required for transcription of the very transcripts that these factors are known to target (Haimovich et al., 2013; Sun et al., 2013). As NMD factors are known to have nuclear functions (Brogna, 1999; Varsally and Brogna, 2012) and have been shown to influence the kinetics of events that occur cotranscriptionally at sites of target gene

transcription (de Turris et al., 2011), it is possible that NMD factors could also have some influence on the transcription rate of native decay targets.

Another recurring theme of NMD-targeting is the observation that NMD pathway genes are often NMD targets themselves. In both *Drosophila* and humans, *Smg5* and *Smg6* are direct NMD targets (Chapin et al., 2014; Huang et al., 2011; Rehwinkel et al., 2005; Yepiskoposyan et al., 2011). In plants, *Smg7* and *Upf3* are targets of the NMD pathway (Rayson et al., 2012). These observations suggest that autoregulation of NMD may be an important aspect of overall NMD function.

Gene ontology and *cis*-acting features of NMD targets

In all organisms, analysis of the ontological categories enriched in genes regulated by NMD encompass a diverse spectrum of functionality, suggesting NMD-mediated regulation could be involved in many cellular processes. In *S. cerevisiae*, types of functional categories of NMD targets include genes involved in telomere maintenance, thiamine biosynthesis, splicing, nitrogen metabolism, and DNA repair . In *S. pombe*, genes that are induced upon nitrogen starvation are also regulated by NMD. In *Drosophila*, NMD appears to regulate an especially diverse array of biological processes including defense response, amino acid metabolism, signal transduction, development, and cell adhesion. In mammals, NMD targets are enriched in genes involved in cell death and amino acid metabolism (McIlwain et al., 2010; Mendell et al., 2004). In plants, NMD targets genes involved in the transcriptional response to pathogens (Rayson et al., 2012).

Taken together, these data suggest that NMD could potentially regulate numerous biological processes; however, certain processes stand out. For instance in *Drosophila*,

the observation that many NMD target genes have known cell cycle functions is consistent with other observation that NMD mutants are defective in cell cycle progression. However, a direct link between specific substrate RNAs and the physiologically relevant functions of NMD has rarely been demonstrated. NMD mutants in *S. pombe* are stress sensitive, and double mutants for NMD genes and the direct target gene *rex2* partially rescue the sensitivity to oxidation observed in NMD single mutants (Matia-González et al., 2013). In *S. cerevisiae*, the magnesium transporter *Alr1* is upregulated in NMD mutants, leading to an increase in intracellular magnesium that decreases the fidelity of translation termination (Johansson and Jacobson, 2010). In *Arabidopsis*, where NMD mutants have been shown to upregulate genes involved in the response to pathogen infection (Rayson et al., 2012), double mutants for *Upf1* and *PAD4*, (a transcription factor critical for the transcriptional response to pathogens, but not a direct NMD target) rescued some of the NMD-dependent defects, indicating that NMD mutants may suffer from an overactive immune response (Riehs-Kearnan et al., 2012). Chapter 3 describes a screen that was designed to identify specific target genes that mediate the critical functions of NMD in *Drosophila*.

Genome-wide analysis of NMD targeting can also be used to test hypotheses about and uncover the *cis*-acting sequence features that allow native genes to be recognized by NMD. For instance, in most organisms, genes with long 3' UTRs and uORFs, both of which have been shown to sensitize transcripts to NMD, are enriched in lists of NMD-targeted genes. When the link between alternative splicing and NMD is examined, many targeted isoforms include PTCs, suggesting that NMD does preferentially target such transcripts. Additionally, genes that contain 3' UTR introns are

targeted in mammals (Mendell et al., 2004). These data demonstrate that genes containing such features are preferentially targeted by NMD. However, there are many genes in all experimental systems that contain no predicted NMD-sensitizing features, yet appear to be directly targeted. For instance, in *S. cerevisiae*, only one-third of targets contain known NMD-stimulating features, such as uORFs or long 3' UTRs. Moreover, many genes that do contain such sensitizing features are not targeted by NMD, suggesting yet undiscovered transcript features that are required to sensitize transcripts to regulation by NMD, or alternatively, sequences that protect transcripts from NMD are present.

NMD and human disease

The study of NMD has often centered around its role in human disease. It is estimated that 11% of all inherited human genetic disorders are due to nonsense mutations (Mort et al., 2008). Therefore, it is possible that NMD is involved in the surveillance of all these disease-associated mRNAs. Supporting this conclusion is the observation that for several well-characterized disorders, polarity of the PTC (generally positioned more 5' or more 3') can drastically change the severity and the Mendelian inheritance pattern (dominant or recessive) of the associated disorder (Bhuvanagiri et al., 2010). These observations are interpreted as a differential sensitivity of the PTC-bearing transcript to NMD, for example, if the PTC is positioned in a way that can be interpreted by the NMD machinery as premature. Examples of this PTC positional polarity include well studied diseases such as spinal muscular atrophy, muscular dystrophy, and cystic fibrosis.

In a few cases, the positional effects of PTC localization have been conclusively shown to be caused by NMD. For example, familial *β*-thalassemia can be inherited in either a dominant or recessive form, both of which can be caused by nonsense mutations in the *β-globin* gene. Alleles positioned towards the 5' end of the gene create a sensitive NMD substrate, and mutant transcripts are degraded. These alleles confer a recessive mode of inheritance as the wild-type copy in heterozygotes compensates for the nonsense allele. However, PTCs positioned in the 3' terminal exon are not detected by NMD, presumably because they lack downstream EJCs, and the mutant transcript is found at wild type levels (Hall and Thein, 1994). These alleles show a dominant mode of inheritance, which stems from a buildup of truncated *β-globin* protein in erythrocytes, and the mutant protein causes a dominant interfering effect on the production of hemoglobin, which overwhelms the cells proteolytic machinery resulting in cell death (Thein et al., 1990).

Because NMD can potentially influence the clinical symptoms of many genetic diseases, there has been considerable effort to devise pharmacological interventions to manipulate the function of the NMD pathway (Durand et al., 2007; Keeling et al., 2013; Kuzmiak and Maquat, 2006; Martin et al., 2014). For instance, if NMD can be inhibited in cases where decay of mutant transcripts eliminates the expression of (partially) functional truncated peptides, it may help to rescue some of the clinical symptoms associated with disease. In a proof-of-principle study, inhibition of NMD was shown to rescue some of the cellular phenotypes associated with mutations in collagen VIα2, which results in the development of Ullrich Disease (Usuki et al., 2006). In this case, pharmacological inhibition of NMD using caffeine or wortmannin (inhibitors of *Smg1*),

or gene silencing of NMD components in patient-derived fibroblasts resulted in rescue of the expression levels of collagen VI. Demonstrating the physiological relevance of this rescue, NMD inhibition also rescued the weakened cell/ECM interactions typical of Ullrich patient fibroblast (Usuki et al., 2006).

Summary

Throughout this thesis, the many genetic techniques available in *Drosophila* are used to study two main questions surrounding NMD: what are the direct native targets of the NMD pathway and what are the functional consequences of this targeting? Chapter 2 describes the development of an NMD reactivation assay for use in whole *Drosophila* aimed at discovering novel *in vivo* direct targets. In Chapter 3, the functional consequences of targeting a potentially important direct target of NMD, *Gadd45*, are analyzed. In Chapter 4, a screen aimed at the identification of tissues where NMD function is most important for viability is described. This screen suggests that NMD function in neuronal tissue may be sufficient for viability in *Drosophila*. Taken together, our data suggest a model where NMD regulates a small number of target genes, including *Gadd45*, in a way that is essential for viability, and that function of NMD in neuronal tissue could be sufficient to rescue the effects of NMD inhibition.

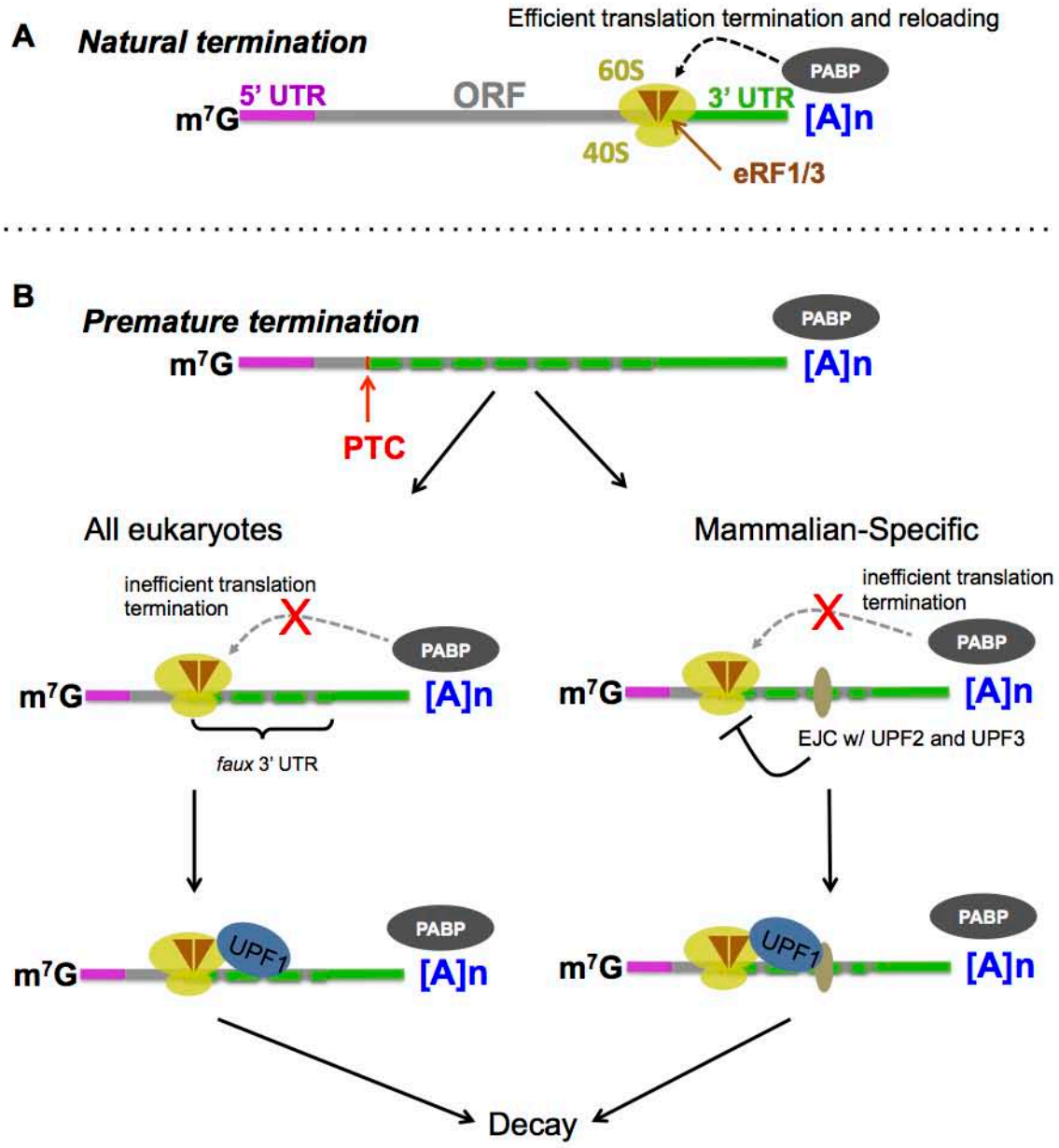


Figure 1.1: Models of NMD targeting: (A) Translation termination that occurs at natural termination codons is efficient. Interactions between PABP and eukaryotic release factors promote efficient release and reloading of ribosomal subunits to the 5' UTR of transcripts. (B) Termination that occurs at premature termination codons is inefficient, due to lack of positive interactions between PABP, and release factors. In all eukaryotes, a long or *faux* 3' UTR has been shown to trigger NMD. In mammals, the presence of an EJC downstream of the termination codon has been shown to sensitize transcripts to NMD.

Table 1.1: Molecular functions and mutant phenotypes of NMD components. PIKK stands for phosphatidylinositol 3-kinase-related kinase; EJC, exon junction complex; E, embryonic day.

		Mutant Phenotype					
Gene	**Function**	**Yeasts**	**C. elegans**	**D. melanogaster**	**Zebrafish**	**Mouse**	**Human**
Upf1	RNA-Helicase	Stress sensitivity	Deformed vulva and bursa	L2 lethal	Inviability w/ impaired brain, eye and somite morphology.	Death before E7.5	N/D
Upf2	EJC adapter					Death before E9.5	N/D
Upf3	EJC component			Viable		N/D	Mental retardation
Smg1	PIKK kinase	N/A			N/D	Death before E12.5	N/D
Smg5	Binds PP2A			Pupal lethal	Inviability w/ impaired brain, eye and somite morphology.	N/D	N/D
Smg6	PIN-domain endonuclease			Viable		N/D	N/D
Smg7	Binds PP2A			N/A		N/D	N/D

References

Addinall, S.G., Downey, M., Yu, M., Zubko, M.K., Dewar, J., Leake, A., Hallinan, J., Shaw, O., James, K., Wilkinson, D.J., et al. (2008). A genome wide suppressor and enhancer analysis of cdc13-1 reveals varied cellular processes influencing telomere capping in Saccharomyces cerevisiae. Genetics *180*, 2251–2266.

Addinall, S.G., Holstein, E.-M., Lawless, C., Yu, M., Chapman, K., Banks, A.P., Ngo, H.-P., Maringele, L., Taschuk, M., Young, A., et al. (2011). Quantitative fitness analysis shows that NMD proteins and many other protein complexes suppress or enhance distinct telomere cap defects. PLoS Genet. *7*, e1001362.

Altamura, N., Groudinsky, O., Dujardin, G., and Slonimski, P.P. (1992). NAM7 nuclear gene encodes a novel member of a family of helicases with a Zn-ligand motif and is involved in mitochondrial functions in Saccharomyces cerevisiae. J. Mol. Biol. *224*, 575–587.

Amrani, N., Ganesan, R., Kervestin, S., Mangus, D.A., Ghosh, S., and Jacobson, A. (2004). A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. Nature *432*, 112–118.

Anders, K.R., Grimson, A., and Anderson, P. (2003). SMG-5, required for C.elegans nonsense-mediated mRNA decay, associates with SMG-2 and protein phosphatase 2A. Embo J. *22*, 641–650.

Arribas-Layton, M., Wu, D., Lykke-Andersen, J., and Song, H. (2013). Structural and functional control of the eukaryotic mRNA decapping machinery. Biochim. Biophys. Acta. *1829*, 580–589.

Avery, P., Vicente-Crespo, M., Francis, D., Nashchekina, O., Alonso, C.R., and Palacios, I.M. (2011). Drosophila Upf1 and Upf2 loss of function inhibits cell growth and causes animal death in a Upf3-independent manner. Rna *17*, 624–638.

Azzalin, C.M., and Lingner, J. (2006). The human RNA surveillance factor UPF1 is required for S phase progression and genome stability. Curr. Biol. *16*, 433–439.

Azzalin, C.M., Reichenbach, P., Khoriauli, L., Giulotto, E., and Lingner, J. (2007). Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. Science *318*, 798–801.

Barbee, S.A., Estes, P.S., Cziko, A.-M., Hillebrand, J., Luedeman, R.A., Coller, J.M., Johnson, N., Howlett, I.C., Geng, C., Ueda, R., et al. (2006). Staufen- and FMRP-containing neuronal RNPs are structurally and functionally related to somatic P bodies. Neuron *52*, 997–1009.

Barberan-Soler, S., Lambert, N.J., and Zahler, A.M. (2009). Global analysis of alternative splicing uncovers developmental regulation of nonsense-mediated decay in C. elegans. Rna *15*, 1652–1660.

Behm-Ansmant, I., Kashima, I., Rehwinkel, J., Saulière, J., Wittkopp, N., and Izaurralde, E. (2007). mRNA quality control: an ancient machinery recognizes and degrades mRNAs with nonsense codons. FEBS Lett. *581*, 2845–2853.

Belasco, J.G. (1993). Control of messenger RNA stability (Waltham: Academic Press).

Ben-Dov, C., Hartmann, B., Lundgren, J., and Valcarcel, J. (2008). Genome-wide analysis of alternative pre-mRNA splicing. J. Biol. Chem. *283*, 1229–1233.

Bhattacharya, A., Czaplinski, K., Trifillis, P., He, F., Jacobson, A., and Peltz, S.W. (2000). Characterization of the biochemical properties of the human Upf1 gene product that is involved in nonsense-mediated mRNA decay. Rna *6*, 1226–1235.

Bhuvanagiri, M., Schlitter, A.M., Hentze, M.W., and Kulozik, A.E. (2010). NMD: RNA biology meets human genetic medicine. Biochem. J. *430*, 365–377.

Brogna, S. (1999). Nonsense mutations in the alcohol dehydrogenase gene of Drosophila melanogaster correlate with an abnormal 3' end processing of the corresponding pre-mRNA. Rna *5*, 562–573.

Bruno, I.G., Karam, R., Huang, L., Bhardwaj, A., Lou, C.H., Shum, E.Y., Song, H.-W., Corbett, M.A., Gifford, W.D., Gécz, J., et al. (2011). Identification of a microRNA that activates gene expression by repressing nonsense-mediated RNA decay. Mol. Cell. *42*, 500–510.

Cali, B.M., Kuchma, S.L., Latham, J., and Anderson, P. (1999). smg-7 is required for mRNA surveillance in Caenorhabditis elegans. Genetics *151*, 605–616.

Carter, M.S., Doskow, J., Morris, P., Li, S., Nhim, R.P., Sandstedt, S., and Wilkinson, M.F. (1995). A regulatory mechanism that detects premature nonsense codons in T-cell receptor transcripts in vivo is reversed by protein synthesis inhibitors in vitro. J. Biol. Chem. *270*, 28995–29003.

Chamieh, H., Ballut, L., Bonneau, F., and Le Hir, H. (2008). NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. Nat. Struct. Mol. Biol. *15*, 85–93.

Chang, Y.-F., Imam, J.S., and Wilkinson, M.F. (2007). The nonsense-mediated decay RNA surveillance pathway. Annu. Rev. Biochem. *76*, 51–74.

Chapin, A., Hu, H., Rynearson, S.G., Hollien, J., Yandell, M., and Metzstein, M.M. (2014). In vivo determination of direct targets of the nonsense mediated decay pathway in Drosophila. G3 *4,* 485–496.

Colak, D., Ji, S.-J., Porse, B.T., and Jaffrey, S.R. (2013). Regulation of axon guidance by compartmentalized nonsense-mediated mRNA decay. Cell *153*, 1252–1265.

Cui, Y., Hagan, K.W., Zhang, S., and Peltz, S.W. (1995). Identification and

characterization of genes that are required for the accelerated degradation of mRNAs containing a premature translational termination codon. Genes. Dev. *9*, 423–436.

Culbertson, M.R., Underbrink, K., and Fink, G. (1980). Frameshift suppression in Saccharomyces cerevisiae. II. Genetic properties of group II suppressors. Genetics *95*, 833–853.

Dahlseid, J.N., Puziss, J., Shirley, R.L., Atkin, A.L., Hieter, P., and Culbertson, M.R. (1998). Accumulation of mRNA coding for the ctf13p kinetochore subunit of Saccharomyces cerevisiae depends on the same factors that promote rapid decay of nonsense mRNAs. Genetics *150*, 1019–1035.

de Turris, V., Nicholson, P., Orozco, R.Z., Singer, R.H., and Muhlemann, O. (2011). Cotranscriptional effect of a premature termination codon revealed by live-cell imaging. Rna *17*, 1–14.

Durand, S., Cougot, N., Mahuteau-Betzer, F., Nguyen, C.-H., Grierson, D.S., Bertrand, E., Tazi, J., and Lejeune, F. (2007). Inhibition of nonsense-mediated mRNA decay (NMD) by a new chemical molecule reveals the dynamic of NMD factors in P-bodies. J. Cell Biol. *178*, 1145–1160.

Eberle, A.B., Lykke-Andersen, S., Mühlemann, O., and Jensen, T.H. (2009). SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. Nat. Struct. Mol. Biol. *16*, 49–55.

Eberle, A.B., Stalder, L., Mathys, H., Orozco, R.Z., and Mühlemann, O. (2008). Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region. Plos Biol. *6*, e92.

Frizzell, K.A., Rynearson, S.G., and Metzstein, M.M. (2012). Drosophila mutants show NMD pathway activity is reduced, but not eliminated, in the absence of Smg6. Rna *18*, 1475–1486.

Garneau, N.L., Wilusz, J., and Wilusz, C.J. (2007). The highways and byways of mRNA decay. Nat. Rev. Mol. Cell. Biol. *8*, 113–126.

Gatfield, D., and Izaurralde, E. (2004). Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in Drosophila. Nature *429*, 575–578.

Gehring, N.H., Lamprinaki, S., Kulozik, A.E., and Hentze, M.W. (2009). Disassembly of exon junction complexes by PYM. Cell *137*, 536–548.

Ghosh, S., Ganesan, R., Amrani, N., and Jacobson, A. (2010). Translational competence of ribosomes released from a premature termination codon is modulated by NMD factors. Rna *16*, 1832–1847.

Giorgi, C., Yeo, G.W., Stone, M.E., Katz, D.B., Burge, C., Turrigiano, G., and Moore, M.J. (2007). The EJC factor eIF4AIII modulates synaptic strength and neuronal protein

expression. Cell *130*, 179–191.

González, C.I., Ruiz-Echevarría, M.J., Vasudevan, S., Henry, M.F., and Peltz, S.W. (2000). The yeast hnRNP-like protein Hrp1/Nab4 marks a transcript for nonsense-mediated mRNA decay. Mol. Cell *5*, 489–499.

Guan, Q., Zheng, W., Tang, S., Liu, X., Zinkel, R.A., Tsui, K.-W., Yandell, B.S., and Culbertson, M.R. (2006). Impact of nonsense-mediated mRNA decay on the global expression profile of budding yeast. PLoS Genet. *2*, e203.

Haimovich, G., Medina, D.A., Causse, S.Z., Garber, M., Millán-Zambrano, G., Barkai, O., Chávez, S., Pérez-Ortín, J.E., Darzacq, X., and Choder, M. (2013). Gene expression is circular: factors for mRNA degradation also foster mRNA synthesis. Cell *153*, 1000–1011.

Hall, G.W., and Thein, S. (1994). Nonsense codon mutations in the terminal exon of the beta-globin gene are not associated with a reduction in beta-mRNA accumulation: a mechanism for the phenotype of dominant beta-thalassemia. Blood *83*, 2031–2037.

Hansen, K.D., Lareau, L.F., Blanchette, M., Green, R.E., Meng, Q., Rehwinkel, J., Gallusser, F.L., Izaurralde, E., Rio, D.C., Dudoit, S., et al. (2009). Genome-wide identification of alternative splice forms down-regulated by nonsense-mediated mRNA decay in Drosophila. PLoS Genet. *5*, e1000525.

He, F., Peltz, S.W., Donahue, J.L., Rosbash, M., and Jacobson, A. (1993). Stabilization and ribosome association of unspliced pre-mRNAs in a yeast upf1- mutant. Proc. Natl. Acad. Sci. USA *90*, 7034–7038.

He, F., Li, X., Spatrick, P., Casillo, R., Dong, S., and Jacobson, A. (2003). Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5“ to 3” mRNA decay pathways in yeast. Mol. Cell *12*, 1439–1452.

Hodgkin, J., Papp, A., Pulak, R., Ambros, V., and Anderson, P. (1989). A new kind of informational suppression in the nematode Caenorhabditis elegans. Genetics *123*, 301–313.

Hogg, J.R., and Goff, S.P. (2010). Upf1 senses 3'UTR length to potentiate mRNA decay. Cell *143*, 379–389.

Houseley, J., LaCava, J., and Tollervey, D. (2006). RNA-quality control by the exosome. Nat. Rev. Mol. Cell Biol. *7*, 529–539.

Huang, L., Lou, C.-H., Chan, W., Shum, E.Y., Shao, A., Stone, E., Karam, R., Song, H.-W., and Wilkinson, M.F. (2011). RNA homeostasis governed by cell type-specific and branched feedback loops acting on NMD. Mol. Cell *43*, 950–961.

Huntzinger, E., Kashima, I., Fauser, M., Saulière, J., and Izaurralde, E. (2008). SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan.

Rna *14*, 2609–2617.

Hurt, J.A., Robertson, A.D., and Burge, C.B. (2013). Global analyses of UPF1 binding and function reveals expanded scope of nonsense-mediated mRNA decay. Genome Res. *23*, 1636–1650.

Hwang, J., and Maquat, L.E. (2011). Nonsense-mediated mRNA decay (NMD) in animal embryogenesis: to die or not to die, that is the question. Curr. Opin. Genet. Dev. *21*, 422–430.

Hwang, J., Sato, H., Tang, Y., Matsuda, D., and Maquat, L.E. (2010). UPF1 association with the cap-binding protein, CBP80, promotes nonsense-mediated mRNA decay at two distinct steps. Mol. Cell *39*, 396–409.

Ivanov, I.P., Simin, K., Letsou, A., Atkins, J.F., and Gesteland, R.F. (1998). The Drosophila gene for antizyme requires ribosomal frameshifting for expression and contains an intronic gene for snRNP Sm D3 on the opposite strand. Mol. Cell Biol. *18*, 1553–1561.

Ivanov, P.V., Gehring, N.H., Kunz, J.B., Hentze, M.W., and Kulozik, A.E. (2008). Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. Embo. J. *27*, 736–747.

Johansson, M.J.O., and Jacobson, A. (2010). Nonsense-mediated mRNA decay maintains translational fidelity by limiting magnesium uptake. Genes Dev. *24*, 1491–1495.

Johansson, M.J.O., He, F., Spatrick, P., Li, C., and Jacobson, A. (2007). Association of yeast Upf1p with direct substrates of the NMD pathway. Proc. Natl. Acad. Sci. USA *104*, 20872–20877.

Johns, L., Grimson, A., Kuchma, S.L., Newman, C.L., and Anderson, P. (2007). Caenorhabditis elegans SMG-2 selectively marks mRNAs containing premature translation termination codons. Mol. Cell Biol. *27*, 5630–5638.

Johnston, L.A. (2009). Competitive interactions between cells: death, growth, and geography. Science *324*, 1679–1682.

Jones, R., Wang, F., Luo, Y., Yu, C., Jin, C., Suzuki, T., Kan, M., and McKeehan, W. (2001). The nonsense-mediated decay pathway and mutually exclusive expression of alternatively spliced FGFR2IIIb and -IIIc mRNAs. J. Biol. Chem. *276*, 4158–4167.

Kashima, I., Yamashita, A., Izumi, N., Kataoka, N., Morishita, R., Hoshino, S., Ohno, M., Dreyfuss, G., and Ohno, S. (2006). Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. Genes Dev. *20*, 355–367.

Kebaara, B.W., and Atkin, A.L. (2009). Long 3'-UTRs target wild-type mRNAs for nonsense-mediated mRNA decay in Saccharomyces cerevisiae. Nucleic Acids Res. *37*,

2771–2778.

Keeling, K.M., Wang, D., Dai, Y., Murugesan, S., Chenna, B., Clark, J., Belakhov, V., Kandasamy, J., Velu, S.E., Baasov, T., et al. (2013). Attenuation of nonsense-mediated mRNA decay enhances in vivo nonsense suppression. PLoS ONE *8*, e60478.

Kervestin, S., and Jacobson, A. (2012). NMD: a multifaceted response to premature translational termination. Nat. Rev. Mol. Cell Biol. *13*, 700–712.

Kervestin, S., Li, C., Buckingham, R., and Jacobson, A. (2012). Testing the faux-UTR model for NMD: Analysis of Upf1p and Pab1p competition for binding to eRF3/Sup35p. Biochimie. *94,* 1560–1571

Kuzmiak, H.A., and Maquat, L.E. (2006). Applying nonsense-mediated mRNA decay research to the clinic: progress and challenges. Trends Mol. Med. *12*, 306–316.

Lareau, L.F., Brooks, A.N., Soergel, D.A.W., Meng, Q., and Brenner, S.E. (2007a). The coupling of alternative splicing and nonsense-mediated mRNA decay. Adv. Exp. Med. Biol. *623*, 190–211.

Lareau, L.F., Inada, M., Green, R.E., Wengrod, J.C., and Brenner, S.E. (2007b). Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. Nature *446*, 926–929.

Le Hir, H., Moore, M.J., and Maquat, L.E. (2000). Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon-exon junctions. Genes Dev. *14*, 1098–1108.

Leeds, P., Peltz, S.W., Jacobson, A., and Culbertson, M.R. (1991). The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon. Genes Dev. *5*, 2303–2314.

Leeds, P., Wood, J.M., Lee, B.S., and Culbertson, M.R. (1992). Gene products that promote mRNA turnover in Saccharomyces cerevisiae. Mol. Cell Biol. *12*, 2165–2177.

Lejeune, F., Li, X., and Maquat, L.E. (2003). Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. Mol. Cell *12*, 675–687.

Lewis, B., Green, R., and Brenner, S.E. (2003). Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proc. Natl. Acad. Sci. USA *100*, 189–192.

Long, A.A., Mahapatra, C.T., Woodruff, E.A., Rohrbough, J., Leung, H.-T., Shino, S., An, L., Doerge, R.W., Metzstein, M.M., Pak, W.L., et al. (2010). The nonsense-mediated decay pathway maintains synapse architecture and synaptic vesicle cycle efficacy. J. Cell Sci. *123*, 3303–3315.

Longman, D., Plasterk, R.H.A., Johnstone, I.L., and Cáceres, J.F. (2007). Mechanistic insights and identification of two novel factors in the C. elegans NMD pathway. Genes Dev. *21*, 1075–1085.

Losson, R., and Lacroute, F. (1979). Interference of nonsense mutations with eukaryotic messenger RNA stability. Proc. Natl. Acad. Sci. USA *76*, 5134–5137.

Maderazo, A.B., He, F., Mangus, D.A., and Jacobson, A. (2000). Upf1p control of nonsense mRNA translation is regulated by Nmd2p and Upf3p. Mol. Cell Biol. *20*, 4591–4603.

Maderazo, A.B., Belk, J.P., He, F., and Jacobson, A. (2003). Nonsense-containing mRNAs that accumulate in the absence of a functional nonsense-mediated mRNA decay pathway are destabilized rapidly upon its restitution. Mol. Cell Biol. *23*, 842–851.

Maquat, L.E., Kinniburgh, A.J., Rachmilewitz, E.A., and Ross, J. (1981). Unstable beta-globin mRNA in mRNA-deficient beta o thalassemia. Cell *27*, 543–553.

Maquat, L.E., Tarn, W.-Y., and Isken, O. (2010). The pioneer round of translation: features and functions. Cell *142*, 368–374.

Marquardt, S., Hazelbaker, D.Z., and Buratowski, S. (2011). Distinct RNA degradation pathways and 3' extensions of yeast non-coding RNA species. Transcription *2*, 145–154.

Martin, L., Grigoryan, A., Wang, D., Wang, J., Breda, L., Rivella, S., Cardozo, T., and Gardner, L.B. (2014). Identification and characterization of small molecules that inhibit nonsense-mediated RNA decay and suppress nonsense p53 mutations. Cancer Res. *74*, 3104–3113.

Matia-González, A.M., Hasan, A., Moe, G.H., Mata, J., and Rodríguez-Gabriel, M.A. (2013). Functional characterization of Upf1 targets in Schizosaccharomyces pombe. RNA Biol. *10*, 1–9.

McGlincy, N., and Smith, C. (2008). Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense? Trends Biochem. Sci. *33*, 385–393.

McIlwain, D.R., Pan, Q., Reilly, P.T., Elia, A.J., McCracken, S., Wakeham, A.C., Itie-Youten, A., Blencowe, B.J., and Mak, T.W. (2010). Smg1 is required for embryogenesis and regulates diverse genes via alternative splicing coupled to nonsense-mediated mRNA decay. Proc. Natl. Acad. Sci. USA *107*, 12186–12191.

Medghalchi, S.M., Frischmeyer, P.A., Mendell, J.T., Kelly, A.G., Lawler, A.M., and Dietz, H.C. (2001). Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. Hum. Mol. Genet. *10*, 99–105.

Mendell, J.T., Medghalchi, S.M., Lake, R.G., Noensie, E.N., and Dietz, H.C. (2000). Novel Upf2p orthologues suggest a functional link between translation initiation and nonsense surveillance complexes. Mol. Cell. Biol. *20*, 8944–8957.

Mendell, J.T., Sharifi, N.A., Meyers, J.L., Martinez-Murillo, F., and Dietz, H.C. (2004). Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nat. Genet. *36*, 1073–1078.

Metzstein, M.M., and Krasnow, M.A. (2006). Functions of the nonsense-mediated mRNA decay pathway in Drosophila development. PLoS Genet. *2*, e180.

Mitchell, P., and Tollervey, D. (2003). An NMD pathway in yeast involving accelerated deadenylation and exosome-mediated 3“-->5” degradation. Mol. Cell *11*, 1405–1413.

Mitrovich, Q.M., and Anderson, P. (2000). Unproductively spliced ribosomal protein mRNAs are natural targets of mRNA surveillance in C. elegans. Genes Dev. *14*, 2173–2184.

Morrison, M., Harris, K.S., and Roth, M.B. (1997). smg mutants affect the expression of alternatively spliced SR protein mRNAs in Caenorhabditis elegans. Proc. Natl. Acad. Sci. USA *94*, 9782–9785.

Mort, M., Ivanov, D., Cooper, D.N., and Chuzhanova, N.A. (2008). A meta-analysis of nonsense mutations causing human genetic disease. Hum. Mutat. *29*, 1037–1047.

Muhlrad, D., and Parker, R. (1999). Aberrant mRNAs with extended 3' UTRs are substrates for rapid degradation by mRNA surveillance. Rna *5*, 1299–1307.

Nagy, E., and Maquat, L.E. (1998). A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. Trends Biochem. Sci. *23*, 198–199.

Neu-Yilik, G., Gehring, N., Hentze, M., and Kulozik, A. (2004). Nonsense-mediated mRNA decay: from vacuum cleaner to Swiss army knife. Genome Biol. *5*, 218.

Ni, J., Grate, L., Donohue, J., Preston, C., Nobida, N., O'Brien, G., Shiue, L., Clark, T., Blume, J., and Ares, M. (2007). Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. Genes Dev. *21*, 708–718.

Okada-Katsuhata, Y., Yamashita, A., Kutsuzawa, K., Izumi, N., Hirahara, F., and Ohno, S. (2012). N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. Nucleic Acids Res. *40*, 1251–1266.

Palacios, I.M. (2013). Nonsense-mediated mRNA decay: from mechanistic insights to impacts on human health. Brief Funct. Genomics *12*, 25–36.

Peltz, S.W., Brown, A.H., and Jacobson, A. (1993). mRNA destabilization triggered by premature translational termination depends on at least three cis-acting sequence elements and one trans-acting factor. Genes Dev. *7*, 1737–1754.

Pisarev, A.V., Skabkin, M.A., Pisareva, V.P., Skabkina, O.V., Rakotondrafara, A.M.,

Hentze, M.W., Hellen, C.U.T., and Pestova, T.V. (2010). The role of ABCE1 in eukaryotic posttermination ribosomal recycling. Mol. Cell *37*, 196–210.

Plath, N., Ohana, O., Dammermann, B., Errington, M.L., Schmitz, D., Gross, C., Mao, X., Engelsberg, A., Mahlke, C., Welzl, H., et al. (2006). Arc/Arg3.1 is essential for the consolidation of synaptic plasticity and memories. Neuron *52*, 437–444.

Ramani, A.K., Nelson, A.C., Kapranov, P., Bell, I., Gingeras, T.R., and Fraser, A.G. (2009). High resolution transcriptome maps for wild-type and nonsense-mediated decay-defective Caenorhabditis elegans. Genome Biol. *10*, R101.

Rayson, S., Arciga-Reyes, L., Wootton, L., De Torres Zabala, M., Truman, W., Graham, N., Grant, M., and Davies, B. (2012). A role for nonsense-mediated mRNA decay in plants: pathogen responses are induced in Arabidopsis thaliana NMD mutants. PLoS ONE *7*, e31917.

Rehwinkel, J., Letunic, I., Raes, J., Bork, P., and Izaurralde, E. (2005). Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. Rna *11*, 1530–1544.

Riehs, N., Akimcheva, S., Puizina, J., Bulankova, P., Idol, R.A., Siroky, J., Schleiffer, A., Schweizer, D., Shippen, D.E., and Riha, K. (2008). Arabidopsis SMG7 protein is required for exit from meiosis. J. Cell Sci. *121*, 2208–2216.

Riehs-Kearnan, N., Gloggnitzer, J., Dekrout, B., Jonak, C., and Riha, K. (2012). Aberrant growth and lethality of Arabidopsis deficient in nonsense-mediated RNA decay factors is caused by autoimmune-like response. Nucleic Acids Res. *40*, 5615–5624.

Rodríguez-Gabriel, M.A., Watt, S., Bähler, J., and Russell, P. (2006). Upf1, an RNA helicase required for nonsense-mediated mRNA decay, modulates the transcriptional response to oxidative stress in fission yeast. Mol. Cell Biol. *26*, 6347–6356.

Sato, H., and Maquat, L.E. (2009). Remodeling of the pioneer translation initiation complex involves translation and the karyopherin importin beta. Genes Dev. *23*, 2537–2550.

Shandilya, J., and Roberts, S.G.E. (2012). The transcription cycle in eukaryotes: from productive initiation to RNA polymerase II recycling. Biochim. Biophys. Acta *1819*, 391–400.

Stowers, R.S., and Schwarz, T.L. (1999). A genetic method for generating Drosophila eyes composed exclusively of mitotic clones of a single genotype. Genetics *152*, 1631–1639.

Sun, M., Schwalb, B., Pirkl, N., Maier, K.C., Schenk, A., Failmezger, H., Tresch, A., and Cramer, P. (2013). Global analysis of eukaryotic mRNA degradation reveals xrn1-dependent buffering of transcript levels. Mol. Cell *52*, 52–62.

Tani, H., Imamachi, N., Salam, K.A., Mizutani, R., Ijiri, K., Irie, T., Yada, T., Suzuki, Y., and Akimitsu, N. (2012a). Identification of hundreds of novel UPF1 target transcripts by direct determination of whole transcriptome stability. RNA Biol. *9*, 1370–1379.

Tani, H., Mizutani, R., Salam, K.A., Tano, K., Ijiri, K., Wakamatsu, A., Isogai, T., Suzuki, Y., and Akimitsu, N. (2012b). Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals. Genome Res. *22*, 947–956.

Tarpey, P.S., Raymond, F.L., Nguyen, L.S., Rodriguez, J., Hackett, A., Vandeleur, L., Smith, R., Shoubridge, C., Edkins, S., Stevens, C., et al. (2007). Mutations in UPF3B, a member of the nonsense-mediated mRNA decay complex, cause syndromic and nonsyndromic mental retardation. Nat. Genet. *39*, 1127–1133.

Thein, S.L., Hesketh, C., Taylor, P., Temperley, I.J., Hutchinson, R.M., Old, J.M., Wood, W.G., Clegg, J.B., and Weatherall, D.J. (1990). Molecular basis for dominantly inherited inclusion body beta-thalassemia. Proc. Natl. Acad. Sci. USA *87*, 3924–3928.

Thompson, D.M., and Parker, R. (2007). Cytoplasmic decay of intergenic transcripts in Saccharomyces cerevisiae. Mol. Cell Biol. *27*, 92–101.

Usuki, F., Yamashita, A., Kashima, I., Higuchi, I., Osame, M., and Ohno, S. (2006). Specific inhibition of nonsense-mediated mRNA decay components, SMG-1 or Upf1, rescues the phenotype of Ullrich disease fibroblasts. Mol. Ther. *14*, 351–360.

Varsally, W., and Brogna, S. (2012). UPF1 involvement in nuclear functions. Biochem. Soc. Trans. *40*, 778–783.

Wang, J., Vock, V.M., Li, S., Olivas, O.R., and Wilkinson, M.F. (2002). A quality control pathway that down-regulates aberrant T-cell receptor (TCR) transcripts by a mechanism requiring UPF2 and translation. J. Biol. Chem. *277*, 18489–18493.

Weischenfeldt, J., Damgaard, I., Bryder, D., Theilgaard-Mönch, K., Thoren, L.A., Nielsen, F.C., Jacobsen, S.E.W., Nerlov, C., and Porse, B.T. (2008). NMD is essential for hematopoietic stem and progenitor cells and for eliminating by-products of programmed DNA rearrangements. Genes Dev. *22*, 1381–1396.

Weischenfeldt, J., Waage, J., Tian, G., Zhao, J., Damgaard, I., Jakobsen, J.S., Kristiansen, K., Krogh, A., Wang, J., and Porse, B.T. (2012). Mammalian tissues defective in nonsense-mediated mRNA decay display highly aberrant splicing patterns. Genome Biol. *13*, R35.

Welch, E.M., and Jacobson, A. (1999). An internal open reading frame triggers nonsense-mediated decay of the yeast SPT10 mRNA. Embo. J. *18*, 6134–6145.

Wittkopp, N., Huntzinger, E., Weiler, C., Saulière, J., Schmidt, S., Sonawane, M., and Izaurralde, E. (2009). Nonsense-mediated mRNA decay effectors are essential for zebrafish embryonic development and survival. Mol. Cell Biol. *29*, 3517–3528.

Yamashita, A., Ohnishi, T., Kashima, I., Taya, Y., and Ohno, S. (2001). Human SMG-1, a novel phosphatidylinositol 3-kinase-related protein kinase, associates with components of the mRNA surveillance complex and is involved in the regulation of nonsense-mediated mRNA decay. Genes Dev. *15*, 2215–2228.

Yamashita, A., Izumi, N., Kashima, I., Ohnishi, T., Saari, B., Katsuhata, Y., Muramatsu, R., Morita, T., Iwamatsu, A., Hachiya, T., et al. (2009). SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay. Genes Dev. *23*, 1091–1105.

Yepiskoposyan, H., Aeschimann, F., Nilsson, D., Okoniewski, M., and Mühlemann, O. (2011). Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. Rna *17*, 2108–2118.

Yoine, M., Nishii, T., and Nakamura, K. (2006). Arabidopsis UPF1 RNA helicase for nonsense-mediated mRNA decay is involved in seed size control and is essential for growth. Plant Cell Physiol. *47*, 572–580.

Zhang, J., Sun, X., Qian, Y., LaDuca, J.P., and Maquat, L.E. (1998). At least one intron is required for the nonsense-mediated decay of triosephosphate isomerase mRNA: a possible link between nuclear splicing and cytoplasmic translation. Mol. Cell Biol. *18*, 5272–5283.

Zünd, D., Gruber, A.R., Zavolan, M., and Mühlemann, O. (2013). Translation-dependent displacement of UPF1 from coding sequences causes its enrichment in 3' UTRs. Nat. Struct. Mol. Biol. *20*, 936–943.

CHAPTER 2

IN VIVO DETERMINATION OF DIRECT TARGETS OF THE NONSENSE-MEDIATED DECAY PATHWAY IN DROSOPHILA

Reprint of Chapin, A., Hu, H., Rynearson, S.G., Hollien, J., Yandell, M., and Metzstein, M.M. (2014). *In Vivo* Determination of Direct Targets of the Nonsense Mediated Decay Pathway in *Drosophila*. G3 (Bethesda).

Reprinted with permission from G3 (Bethesda)

INVESTIGATION

In Vivo Determination of Direct Targets of the Nonsense-Mediated Decay Pathway in *Drosophila*

Alex Chapin,*,[1] Hao Hu,*,[1,2] Shawn G. Rynearson,* Julie Hollien,[†] Mark Yandell,* and Mark M. Metzstein*,[3]

*Department of Human Genetics and [†]Department of Biology, University of Utah, Salt Lake City, Utah 84112

ABSTRACT Nonsense-mediated messenger RNA (mRNA) decay (NMD) is a mRNA degradation pathway that regulates a significant portion of the transcriptome. The expression levels of numerous genes are known to be altered in NMD mutants, but it is not known which of these transcripts is a direct pathway target. Here, we present the first genome-wide analysis of direct NMD targeting in an intact animal. By using rapid reactivation of the NMD pathway in a *Drosophila melanogaster* NMD mutant and globally monitoring of changes in mRNA expression levels, we can distinguish between primary and secondary effects of NMD on gene expression. Using this procedure, we identified 168 candidate direct NMD targets *in vivo*. Remarkably, we found that 81% of direct target genes do not show increased expression levels in an NMD mutant, presumably due to feedback regulation. Because most previous studies have used up-regulation of mRNA expression as the only means to identify NMD-regulated transcripts, our results provide new directions for understanding the roles of the NMD pathway in endogenous gene regulation during animal development and physiology. For instance, we show clearly that direct target genes have longer 3′ untranslated regions compared with non-targets, suggesting long 3′ untranslated regions target mRNAs for NMD *in vivo*. In addition, we investigated the role of NMD in suppressing transcriptional noise and found that although the transposable element *Copia* is up-regulated in NMD mutants, this effect appears to be indirect.

KEYWORDS
Upf2
reactivation
NMD
Drosophila
RNA-seq

Steady-state messenger RNA (mRNA) expression levels are controlled by a balance of *de novo* transcription and transcript degradation. Nonsense-mediated mRNA decay (NMD) is a degradation mechanism that functions to target mRNAs in a translation-dependent manner (Kervestin and Jacobson 2012). The machinery required to execute NMD is evolutionarily conserved (Behm-Ansmant *et al.* 2007b), with three core proteins, Upf1, Upf2, and Upf3, found throughout the eukaryotes. Four other well-characterized NMD factors, Smg1, Smg6, and the paralogous proteins Smg5 and Smg7, are thought to function as regulators of the core components and are required for NMD in many, but not all organisms (Behm-Ansmant *et al.* 2007b; Riehs *et al.* 2008; Lloyd and Davies 2013). NMD function is required for viability in complex organisms, including *Drosophila* (Metzstein and Krasnow 2006), zebrafish (Wittkopp *et al.* 2009), mammals (Medghalchi *et al.* 2001; McIlwain *et al.* 2010; Weischenfeldt *et al.* 2012), and plants (Kerényi *et al.* 2008). In contrast, in simpler organisms, including *Saccharomyces cerevisiae* (Leeds *et al.* 1991), *Schizosaccharomyces pombe* (Mendell *et al.* 2000), and *Caenorhabditis elegans* (Hodgkin *et al.* 1989), NMD is not required for viability. However, loss of NMD pathway function leads to stress sensitivity in yeasts (Leeds *et al.* 1991; Rodríguez-Gabriel *et al.* 2006) and abnormal morphogenesis in *C. elegans* (Hodgkin *et al.* 1989), indicating that NMD does have important biological roles in these organisms. Loss-of-function mutations in different NMD genes lead to a similar spectrum of defects within an organism (Hodgkin *et al.* 1989; Rehwinkel *et al.* 2005; Metzstein and Krasnow 2006; Wittkopp *et al.* 2009; Avery *et al.* 2011; Frizzell *et al.* 2012), suggesting that, although a number of NMD pathway components may function in NMD-independent processes

doi: 10.1534/g3.113.009357
Manuscript received November 8, 2013; accepted for publication January 14, 2014; published Early Online January 15, 2014.

Sequence data from this article have been deposited with the Sequence Read Archive under nos. SRR896609, SRR896616, SRR503415, SRR503416.
Supporting information is available online at http://www.g3journal.org/lookup/suppl/doi:10.1534/g3.113.009357/-/DC1
Arrays have been submitted to the GEO database at NCBI as series GSE47979.
[1]These authors contributed equally to this work.
[2]Present address: Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030.
[3]Corresponding author: 15 N 2030 E, Salt Lake City, UT 84112.
E-mail: markm@genetics.utah.edu

(Azzalin and Lingner 2006; Roberts *et al.* 2013), NMD is in itself required for proper development and physiology.

It is not known why organisms rely on a functional NMD pathway for viability. The best-known targets of the NMD pathway are transcripts harboring premature termination codons (PTCs). One possibility is that all organisms contain a background load of PTC-containing transcripts, arising from inheritance or by sporadic errors in transcription or splicing, and that the NMD pathway is required to eliminate these erroneous mRNAs to maintain normal cellular function (Neu-Yilik *et al.* 2004). An alternative model is that NMD regulates a set of endogenous, error-free transcripts and this regulation is required for development and viability. One observation in support of this latter model is that loss of NMD pathway function results in changes in the expression levels of many endogenous genes (He *et al.* 2003; Mendell *et al.* 2004; Rehwinkel *et al.* 2005; Metzstein and Krasnow 2006; Weischenfeldt *et al.* 2008; Ramani *et al.* 2009; Wittkopp *et al.* 2009; Yepiskoposyan *et al.* 2011; Rayson *et al.* 2012). These studies have found that up to 15% of all genes in diverse organisms may be under constitutive negative regulation by the NMD pathway. On the basis of these analyses, a model has emerged wherein NMD-mutant phenotypes, including lethality, are due to misregulation of specific native targets (Hwang and Maquat 2011; Palacios 2013). However, it is unclear how many genes that are misregulated in NMD mutants are directly targeted by the NMD pathway *vs.* those genes that are indirectly targeted, and which direct NMD targets are responsible for the organismal defects observed in NMD mutants. Distinguishing between direct and indirect targets is thus important to understanding NMD mutant phenotypes, as well as the mechanisms by which the NMD machinery recognizes endogenous substrates.

In a few studies, investigators have sought to identify direct NMD targets by using a number of different approaches. Splicing-sensitive microarrays have been used to identify instances of differential change in expression of specific mRNA isoforms when NMD is inhibited (Barberan-Soler *et al.* 2009; Hansen *et al.* 2009). Because changes in transcription after NMD disruption are expected to affect all splice isoforms equally, differential NMD-sensitivity of different splice isoforms indicates direct targeting. However, NMD has been shown to functionally regulate splicing factors (Barberan-Soler *et al.* 2009; Weischenfeldt *et al.* 2012), so relative changes in isoform expression may not only be due to direct targeting. Another approach to assess direct NMD targeting is to identify which RNAs are bound to the critical NMD factor Upf1, because direct NMD targets are known to show enriched Upf1 binding (Johansson *et al.* 2007; Hurt *et al.* 2013; Matia-González *et al.* 2013). Such binding does not, however, unequivocally identify NMD substrates, as Upf1 also binds many non-NMD targets (Hogg and Goff 2010; Zünd *et al.* 2013).

Another genome-wide technique that can be used to identify NMD targets is analysis of mRNA decay rates, with the assumption that direct NMD targets will show increased stability upon inhibition of NMD. Using such an approach, Guan *et al.* (2006) compared rates of mRNA decay between *Upf1*$^{+}$ and *Upf1*Δ yeast strains after treatment with the transcriptional inhibitor thiolutin. This analysis revealed that 220 of 616 genes (36%) that are up-regulated in *Upf1*Δ mutants had slower decay than that observed in the *Upf1*$^{+}$ background. Recently, Tani *et al.* (2012a) globally measured mRNA decay rates in HeLa cells depleted of *Upf1* by siRNA knockdown. In this case, the authors used the RNA sequencing-based BRIC-seq technique (Tani *et al.* 2012b), which allows for monitoring of decay without the use of transcriptional inhibitors. They found that the mRNAs of 76 of 324 genes (23%) that were significantly up-regulated after *Upf1* knockdown were stabilized in the absence of *Upf1*. Thus, both these studies conclude that the majority of genes up-regulated in the absence of a functional NMD pathway are not directly targeted by the NMD machinery.

A final approach that has been used to identify direct NMD-pathway substrates is the use of an NMD "reactivation" assay (Maderazo *et al.* 2003; Johansson *et al.* 2007). This assay, developed in yeast cells, uses an inducible system to rapidly restore levels of an NMD gene in a cognate NMD mutant background. Before reactivation, levels of both direct and indirect NMD-regulated genes accumulate. After reactivation, direct targets are expected to be rapidly degraded whereas levels of secondary targets will decrease only after direct target levels are normalized. Thus, by monitoring mRNA levels over a reactivation time course, direct and secondary NMD targets can be distinguished. Using this approach in *S. cerevisiae*, Johansson *et al.* (2007) found that 427 of 792 (54%) of genes that are overexpressed in an NMD mutant appear to be direct targets, as defined by the reactivation assay, and conversely, 68% of direct targets are overexpressed in NMD mutants.

No study has yet been performed to experimentally distinguish direct from indirect NMD targets in an intact, multicellular organism. Here, we describe the use of larval *Drosophila melanogaster* to characterize NMD targeting in such a setting. First, we used RNA sequencing to identify transcripts with altered expression levels in animals mutant for the NMD gene *Upf2*. Next, to identify direct targets, we adapted the yeast NMD reactivation assay so as to distinguish rapidly degrading direct targets from slower responding indirect targets. To this end, we reintroduced *Upf2* into *Upf2*-deficient larvae and identified transcripts that were depleted rapidly on a genome-wide scale. We find that a minority of genes up-regulated in an NMD mutant appear to be direct targets. Strikingly, we also find that a majority of directly targeted genes do not show increased expression at steady state when the NMD pathway is inactivated. Bioinformatic analysis of our candidate direct targets reveals that these genes have on average significantly longer 3′ untranslated regions (UTRs) then nontarget control transcripts, suggesting that a long 3′ UTR is a primary mechanism of NMD targeting in intact *Drosophila*, as has been proposed based on cell culture experiments (Behm-Ansmant *et al.* 2007a). We also use our reactivation assay to demonstrate that an observed increase in retrotransposon expression, known to occur in NMD mutants, is likely due to indirect effects of NMD pathway disruption.

MATERIALS AND METHODS

Fly stocks and genetics

All fly stocks were reared according to standard protocols (Sullivan *et al.* 2000). Balancers used were *FM7i*, *Act5C:GFP*, and *FM7c*. The mutant NMD alleles are described in Metzstein and Krasnow (2006). *Upf2*^{25G} is on the chromosome *y w Upf2*^{25G} *FRT*^{19A} and for a control, we used *y w FRT*^{19A} (Xu and Rubin 1993). The heat-shock GAL4 is *w**; *P{w[+mC]=GAL4-Hsp70.PB}89-2-1* (Bloomington *Drosophila* Stock Center #1799) (Roman *et al.* 2000).

Construction of the *UAS-Upf2* transgene

The long *Upf2* coding sequence plus 3′ UTR (4767 bp) required a multistep procedure to generate the full-length rescuing transgene. First, we used the primer pairs Upf2xF1 and Upf2xR1 to amplify a 140 bp 5′ fragment and the primer pair Upf2xF2 and Upf2xR2 to amplify a 500 bp 3′ fragment at the end of the *Upf2* 3′ UTR, using the full-length *Upf2* cDNA RE04053 (Stapleton *et al.* 2002) as a template. The two fragments were then joined (based on overlapping sequences present in Upf2xR1 and Upf2xF2) using Upf2xF1 and

Upf2xR2. The resulting product was TOPO-TA cloned in pCR4 (Invitrogen, Carlsbad, CA) and sequenced to confirm that no errors were introduced during polymerase chain reaction (PCR) amplification. We then replaced the middle section of this clone (using naturally occurring *Nhe*I and *Bst*BI restriction sites present in the *Upf2* gene) with a 4.3-kb *Nhe*I/*Bst*BI fragment from RE04053. Finally, we subcloned this full-length cDNA using *Not*I and *Sma*I sites present in the primers, into *Not*I/*Stu*I-digested pUAST (Brand and Perrimon 1993). Note, this procedure removes the SV40 3′ UTR present in pUAST, which is itself sensitive to NMD-mediated degradation (Metzstein and Krasnow 2006). Flies were transformed with the pUAST construct using standard *P*-element−mediated transgenesis. Transgenic injections were performed by BestGene Inc. (Chino Hills, CA). Primer sequences listed in Table S4.

RNA isolation and cDNA preparation

Whole animals or cells were homogenized in Trizol (Invitrogen) and total RNA extracted using a standard three-phase chloroform technique (Chomczynski and Sacchi 1987). RNA was purified and concentrated using the RNeasy column (QIAGEN Inc., Gaithersburg, MD), including an on-column treatment with DNase I. Random decamer-primed cDNA was prepared using M-MLV reverse transcriptase (RNase H+) as part of the RETROscript kit (Ambion, Austin, TX). For each reverse transcription reaction, 10 μg of input RNA was used to aid in comparisons of absolute abundance between samples.

RNA sequencing

For each replicate (two control and two mutant), we collected 11 male larvae of genotype *y w* FRT^{19A}/*Y* (control) or *y w* $Upf2^{25G}$ /*Y* (mutant) in a 0- to 4-hr time period past the L2/L3 molt. Male larvae where identified by sorting using the *FM7i*, *Act5C:GFP* balancer (for $Upf2^{25G}$), or by morphological criteria (for the control). Total RNA was isolated using Trizol/chloroform and quality assessed using a 2100 Bioanalyzer (Agilent, Santa Clara, CA). cDNA synthesis and sequencing was performed by Cofactor Genomics (St. Louis, MO). In brief, rRNA was removed from the sample using RiboMinus Eukaryote Kit (Invitrogen), 1 μg of this purified RNA was fragmented using reagents in the Illumina RNA-seq kit (Illumina, San Diego, CA), then reverse transcribed using random hexamers and Superscript II (Invitrogen, Carlsbad, CA), and the second strand synthesized with DNA Pol I and RNase H. To construct a sequencing DNA library, the double-stranded cDNA was blunted, tailed with an A base, ligated with paired-end adaptors in the Illumina RNA-seq kit (Illumina, San Diego, CA), and size selected on a polyacrylamide gel. Sequencing was performed on an Illumina GAIIx platform, according to manufacturers protocols.

Quantification of RNA-seq data

We used the USeq package to identify differentially expressed genes (Nix *et al.* 2008). In brief, USeq counts the number of reads aligned to each annotated gene and uses the DESeq algorithm to prioritize genes according to their *P*-values of being differentially expressed (Anders and Huber 2010). Internally, a negative binomial distribution is adopted by DESeq to model the discrete read counts in each sample and account for the overdispersion of the coverage data.

Reactivation and microarray experiments

RNA was isolated, as described previously, from the following genotypes: $Upf2^{25G}$/*FM7i*, *Act5C:GFP*; *UAS-Upf2/hsp70-GAL4* (experimental), and $Upf2^{25G}$/*FM7i*, *Act5C:GFP*; *hsp70-GAL4/+* (control). Library construction and hybridization were performed at the GNomEx core at the Huntsman Cancer Institute at the University of Utah. cDNA was generated in a poly dT-primed reverse-transcription (RT)-PCR using MMLV-RT. Labeled cRNA was generated by incorporating cyanine 3-CTP or cyanine 5-CTP into a T7 RNA polymerase reaction using the cDNA as template. Hybridization was performed on an Agilent *Drosophila* 44K array using a Agilent SureHyb hybridization oven. Hybridized chips were scanned using a Agilent G2505C Microarray Scanner and signal processing accomplished with Agilent Feature Extraction software, v 10.5. Stepwise normalization was performed with the Limma package in R (Smyth and Speed 2003), including background signal correction, normalization within microarrays, and normalization between microarrays.

Identification of reactivation targets

To identify high confidence reactivation targets, we first applied a number of filtering steps to the data obtained from our microarray time course. We first eliminated genes from the analysis that displayed expression changes of more than 20% from the preheat shock time point to the 4-hr time point in control experiment (without cDNA). This was found to be primarily due to low absolute expression levels. Next, we verified that leaky expression from the *Upf2*:cDNA transgene did not lead to rescue of NMD target gene levels by comparing expression between the experimental and control samples at the prior-to-heat shock time point, and removed any genes that differed in expression by more than 50% between the two condition. In this analysis, we did not observe consistent rescue of target genes. Finally, those genes which did not have high-confidence annotations in FlyBase were also removed.

We then used criteria similar to those used by Johansson *et al.* (2007) to identity reactivation targets in the filtered set. Our primary criterion was a decrease in expression of at least 1.8-fold over the 4-hr time course in the animals carrying the rescue cDNA and that this decrease in expression differed significantly from control animals, which were heat shocked but did not carry the rescuing cDNA. To perform this latter analysis, we first converted each of the four time points in the experiment to a nominal ordinal value: the time point before the heat shock is 1; the time point right after the heat shock is 2; and so on. For each gene represented on the array, we calculated the Pearson correlation coefficient between its expression level at each time point and the ordinal time value. Permutation tests were used to evaluate the significance of association between the expression and the time of treatment. Genes in this regression analysis that had negative slopes and differed significantly from control ($P < 0.1$), as well as the 1.8-fold decrease from the preheat shock time point, were selected as reactivation targets. Genes that did not fulfill either of these criteria were defined as our control, nondirect target set.

Statistical analysis

We performed logistic regressions to evaluate the relationship between NMD targeting and several features by using the reactivation target set and nonreactivation target set. Tested transcript features included predictor variables: 3′ UTR length; 3′ UTR RNA structure, as evaluated by CentroidFold RNA folding energy (Hamada *et al.* 2009); the presence of introns in 3′ UTR; potential stop-codon read-through events, as defined by (Jungreis *et al.* 2011); the length ratio between 3′ UTR and the coding sequence; the presence of bi/poly-cistronic transcripts; and the presence of stop codons in the 3′ UTR. For qualitative features, we set nominal values for the presence or absence of the feature.

Amplicon design and quantitative RT-PCR (qRT-PCR)

Primer3 (http://frodo.wi.mit.edu/) was used to design amplicons for qRT-PCR experiments using a primer size of 23 bp, a melting temperature of 62°, and designed to amplify a region of 60−100 bps. When possible, the amplicon, or the primers themselves, was designed to span splice junctions to minimize the amplification of contaminating genomic DNA. Primer pairs were tested for amplification efficiency with a wild-type cDNA dilution series. The following primers were used: *Gadd45* (1: qGadd45_F1 and qGadd45_R1; 2: qGadd45_F2 and qGadd45_R2); *RP49*: qRP49_F and qRP49_R; *Copia*; qCopia_F and qCopia_R; and *tra*: qtra_F and qtra_R. All qRT-PCR analysis was conducted using iQ SYBR Green master mix and a MyiQ thermal cycler running iQ5 v2.0 software (Bio-Rad, Hercules, CA). *RP49* (*RpL32*) served as a reference gene in all cases. Primer sequences listed in Table S4.

S2 cell culture and RNA interference (RNAi)

We cultured *Drosophila* S2 cells at room temperature in Schneider's media (Invitrogen) supplemented with 10% fetal bovine serum and antibiotics. To deplete cells of Upf1, we amplified a 771 base pair region from the coding sequence of *Upf1* using primers with T7 RNA polymerase sites at the 5′ ends (Upf1_RNAi_F and Upf1_RNAi_R), and used this product to synthesize double-stranded RNA (Megascript kit; Ambion). We then incubated S2 cells with this dsRNA in serum-free media for 45 min, replaced the serum, and allowed the cells to recover for 3 d. We retreated with dsRNA to maintain the knockdown, allowed a 3-d recovery, and then mock-treated or treated the cells with actinomycin D for 1 hr. RNA was collected and purified using Trizol reagent (Invitrogen) and used to synthesize cDNA for qRT-PCR analysis.

RESULTS

Characterization of *Upf2^{25G}* mutant larvae

Previously, we have shown that hemizygous males harboring a strong hypomorphic mutation in *Upf2*, called *25G*, leads to stabilization of NMD-sensitive reporter mRNAs, PTC bearing mutant transcripts, and endogenous targets of NMD (Metzstein and Krasnow 2006). However, we found that, unlike null alleles of *Upf1* or *Upf2*, which impart lethality in the second larval stage (L2), the development of *Upf2^{25G}/Y* animals proceeds normally through these stages (Figure 1A). The start of the third larval stage (L3) also begins normally in *Upf2^{25G}/Y*, but development then slows such that progression into the pupal stage is delayed by approximately 24 hr (Figure 1B). Most *Upf2^{25G}/Y* die as pupae, with only 14% of animals eclosing into adulthood (Figure 1A). Using the *25G* hypomorphic mutation at the early L3 stage thus allows us to examine the effects of NMD pathway disruption without the complications of lethality and developmental delay inherent to *Upf2* null mutants.

Analysis of the transcriptome of *Upf2^{25G}* mutant larvae

To analyze the effect of NMD pathway disruption on endogenous gene expression, we performed deep-sequencing on total RNA extracted from *Upf2^{+}/Y* and *Upf2^{25G}/Y* L3 larvae. Control and mutant males were stage matched by rearing to 0−4 hr after the L2-L3 larval molt, and total RNA was extracted and sequenced on the GAx II Illumina platform (see *Materials and Methods*). Sequencing reads were aligned to the reference genome (dm3) using Tophat software under the pair-end mode (Trapnell *et al.* 2012). We analyzed two biological replicates each of mutant and control samples and were able to align 24−53 million read pairs to the genome per sample (Supporting Information, Table S1). Although we found reads representing expression of 14,699 genes, only 10,434 were detected in all four samples and so this subset was used for subsequent statistical analysis (File S1). Of these 10,434 genes, we identified 413 (4.0%) as being differentially up-regulated at a significance of $P < 0.01$ (Figure 1, C and D; top genes listed in Table S2 and all genes listed in File S2). These genes were up-regulated between 1.8- and 330-fold in *Upf2^{25G}* compared with control samples. A total of 247 genes were found to be down-regulated in *Upf2^{25G}/Y* (Table S2 and File S2). We found that genes encompassing a broad range of absolute expression levels are subjected to NMD-mediated regulation (Figure 1D, Table S2, and File S2).

Identification of NMD reactivation targets

To identify which transcripts are directly targeted by the NMD pathway, we used a pathway reactivation assay. Starting with an NMD mutant, in which expression levels of both direct and indirect targets should be elevated, a wild-type copy of the mutated factor is expressed, and RNA expression levels are measured over time. It is expected that expression levels of transcripts that are direct targets of NMD should decrease rapidly. Conversely, the levels of indirect targets will remain high until direct targets return to wild-type levels. To perform this experiment, we used the bipartite *UAS/GAL4* expression system (Brand and Perrimon 1993) to rapidly restore wild-type *Upf2* from a *UAS:Upf2* transgene in *Upf2^{25G}/Y* mutant L3 larvae by using the *hsp70* enhancer to drive GAL4 in a heat-inducible manner (Figure 2A). We confirmed that the *UAS:Upf2* transgene was functional by testing for rescue; we found that *UAS:Upf2* driven by *Act5C:GAL4* could rescue the null allele *Upf2^{14J}* to full viability in hemizygous males (data not shown). In addition, we observed no lethality or other defects from driving high-level expression of *Upf2*, either constitutively, using *Act5C:GAL4*, or acutely after heat shock with *hsp70:GAL4*, suggesting that these experimental conditions are unlikely to result in ectopic degradation of non-NMD targets.

Our time course consisted of a preheat-shock time point, a t = 0 time point collected directly after a 30-min heat shock, and two further time points, 2 and 4 hr after the completion of the heat shock. As a control for the effect of the heat shock itself on gene expression, we analyzed larvae that harbored *Upf2^{25G}* and heat-shock GAL4 but not the *UAS:Upf2* transgene. These larvae were treated to the same heat-shock regime as the experimental samples. At each of the time points, for both experimental and control samples, RNA was collected and expression levels were measured using microarrays (see *Materials and Methods*). Analysis of *Upf2* expression levels revealed that *Upf2* was activated over our time course approximately 12-fold in experimental animals, but remained unchanged in control animals, as expected (Figure S1).

We next developed a statistical model to test whether, for each gene, reactivation of the NMD pathway led to a change in mRNA abundance. Our primary assumption was that the expression levels of mRNAs that are direct targets of the NMD pathway would decrease rapidly during the course of the experiment. We defined two parameters to define whether a gene was a direct NMD target. First, the expression level at the end of the 4-hr time course had to decrease by 1.8-fold compared with preheat shock levels. Second, changes in expression were compared with control samples (heat shocked, but lacking the *UAS:Upf2* transgene) to control for effects of heat shock and culturing conditions, with a negative binomial model used to compare differences in gene expression at each time point (see *Materials and Methods*). The difference in stability between experimental

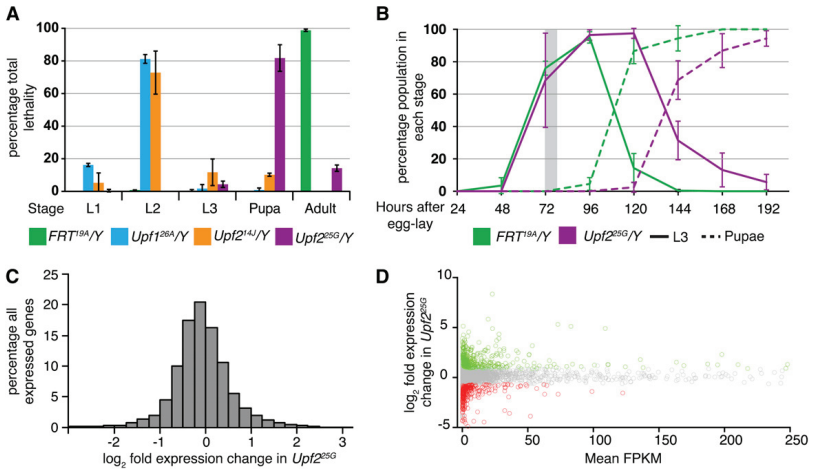


Figure 1 Phenotypic and transcriptome analysis of *Upf2*^{25G}. (A) Effective lethal phase of NMD mutants. Null mutations in *Upf1* (*26A*; cyan) and *Upf2* (*14J*; orange) result in death during the second larval instar, as compared with control (green), the great majority of which survive to adulthood. Animals hemizygous for a hypomorphic allele of *Upf2* (*25G*; purple) die mostly as pupae, with approximately 14% escaping into adulthood. (B) *Upf2*^{25G} */Y* animals develop at a normal rate into L3 but are delayed in this stage of larval development. Solid lines represent the proportion of animals in the L3 stage at each time point; dashed lines represent the proportion in pupal stages. The gray box indicates the collection window (0- to 4–hr-old L3s) in which we compared gene expression changes between control and *Upf2*^{25G}*/Y*. (C) Proportion of genes based on relative expression in *Upf2*^{25G}*/Y* compared with control, based on normalized RNA-seq read depth. (D) Scatter plot of all genes (gray circles) with their average expression level on the x-axis and their relative expression change in *Upf2*^{25G} mutants on the y-axis. Significantly up-regulated and down-regulated genes ($P < 0.01$) are represented by green and red circles, respectively. In (A) and (B), error bars represent ± 1 SD ($n > 118$ for all genotypes).

and control conditions is also described by a best-fit linear regression slope of the log-transformed expression level throughout the time course, which assumes exponential decay of target genes. We refer to this difference in stability as "relative slope" (Figure 2B).

After eliminating transcripts that displayed inconsistent expression levels throughout the time course, such as those that were not expressed at high enough levels for meaningful analysis (see *Materials and Methods* for details), we were left with 6956 genes (File S3). We found a small number of genes (154) that displayed the rapid decrease in expression levels in experimental samples compared to controls defined above, and are thus putative direct targets of the NMD pathway (File S4). We refer to these genes as reactivation targets. The other 6802 genes are defined as nonreactivation targets and are not likely to be directly regulated by the NMD pathway, although we cannot rule out the possibility that some of these are indeed targets but their expression does not change significantly enough to be identified under our experimental conditions.

Among the reactivation targets, we found well-known direct NMD targets, including the NMD pathway components *Smg5* and *Smg6*, whose homologs are direct NMD targets in mammals (Huang *et al.* 2011; Yepiskoposyan *et al.* 2011), and have also been shown to be direct targets in *Drosophila* cell culture (Rehwinkel *et al.* 2005). We also found as a target *Gadd45*, homologs of which are known to be direct NMD targets in mammals (Viegas *et al.* 2007; Huang *et al.* 2011; Tani *et al.* 2012a). To validate that other reactivation targets represent direct substrates of the NMD pathway, we directly measured transcript stability of several of them in *Drosophila* S2 cells with and without a functional NMD pathway. Depletion of *Upf1* by RNAi resulted in stabilization of these mRNAs (Figure S2), indicating our reactivation targets likely represent genuine direct targets of NMD. No particular gene ontology category was enriched in the reactivation data set, implying that NMD may regulate a functionally diverse array of target genes.

Comparison of mutant expression changes and reactivation targets

By combining the expression data from the RNA-seq experiment with our reactivation experiment, we assigned genes to one of four classes (Figure 2B). Some genes, such as *Gadd45*, are up-regulated in *Upf2*^{25G} and were identified as reactivation targets. Others, such as *GstD2*, are up-regulated in mutant animals but did not rapidly decay after reactivation of *Upf2*. Conversely, some genes, such as *Rip11*, are not over-expressed in *Upf2*^{25G}, but still had mRNA levels that decreased rapidly upon *Upf2* reactivation. Finally, most genes are neither up-regulated in *Upf2*^{25G} nor change during the reactivation experiment. This latter class includes genes such as *Act5C* that are known be unaffected by NMD in numerous experiments.

To compare the sets of genes with expression changes in *Upf2*^{25G} mutants with reactivation targets, we restricted our analysis to genes that were both represented on our microarray, present in our deep sequencing data set, and were well annotated in FlyBase (see *Materials and Methods*). A total of 5539 genes fulfilled these criteria (File S5), of which 149 are up-regulated in *Upf2*^{25G} and 125 are reactivation targets. Common to both the up-regulated and reactivation target sets

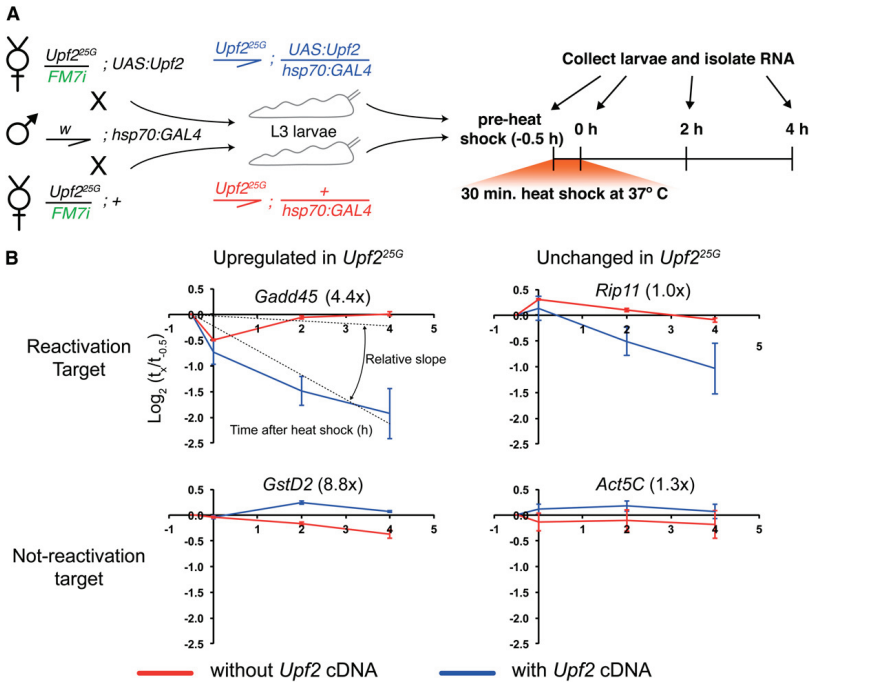


Figure 2 Identification of direct targets of NMD using reactivation of *Upf2*. (A) Crosses used to generate larvae. *Upf2^{25G}/FM7i ; UAS:Upf2* or *Upf2^{25G}/FM7i* females were crossed to *w/Y; hsp70:GAL4* males. *Upf2^{25G}/Y ; UAS:Upf2/ hsp70:GAL4* (experimental, blue) or *Upf2^{25G}/Y ; +/ hsp70:GAL4* (control, red) L3 male larvae were collected based on *Act5C:GFP* carried on the *FM7i* balancer (represented in green) and male gonadal morphology. Larvae of the appropriate genotype were subjected to the heat-shock regime indicated on the right. (B) Examples of genes in the four classes identified; *Gadd45* (up-regulated in *Upf2^{25G}/Y*, reactivation target; upper left); *Rip11* (not up-regulated in *Upf2^{25G}/Y*, reactivation target; upper right); *GstD2* (up-regulated, nonreactivation target; lower left); *Actin5C* (not up-regulated, nonreactivation target). Numbers in parentheses represent the fold change observed in *Upf2^{25G}/Y* compared with control. X-axis represents the time points collected as in (A). Y-axis is relative expression level (on a $\log_2$ scale) compared to the pre-heat shock time point (−0.5 hr) within each genotype. We also define a relative decay slope for each condition, calculated from the regression line of log-transformed data throughout the time course, equal to the slope of the experimental condition (with cDNA) minus the slope of the control condition (without cDNA). Error bars represent ± 1 SD.

were 24 genes. Therefore, of the 149 up-regulated genes, we found only 16% (24/149) are reactivation targets (Figure 3A and Table 1). These data suggest that 84% of genes up-regulated in the *Upf2^{25G}* mutant are indirect targets of the NMD pathway, likely responding to expression changes of direct targets. However, that 16% of genes up-regulated in NMD mutants are found to be reactivation targets represents a significant enrichment over the entire analyzed gene set, in which only 2.3% (125 of 5539) of genes are reactivation targets (Fisher exact test P-value = 5.85×10^{-12}). In addition, the relative slope of the 149 up-regulated genes in *Upf2^{25G}* mutants is skewed negatively, corresponding to degradation during NMD reactivation (Figure 3B, gray bars), compared with the distribution across all 5,539 genes in the analysis (Figure 3B, black bars), also indicating that up-regulated genes are enriched for direct targets.

Most remarkably, these data indicate that the majority (125 of 149) of reactivation targets are not increased at steady state in *Upf2^{25G}* (Figure 3A). Thus, even though these genes appear to be direct targets, their steady expression levels are unaffected by loss of NMD pathway function.

Structural features of genes regulated by the NMD pathway

Several kinds of sequence and structural features are thought to influence whether an mRNA is targeted by NMD (Schweingruber *et al.* 2013). Current models of NMD, particularly in nonvertebrates, suggest 3′ UTR length is a primary signal that stimulates NMD targeting (Kebaara and Atkin 2009; Kervestin and Jacobson 2012). Other features that may direct transcripts for degradation include a weakly structured 3′ UTR (Eberle *et al.* 2008); a low ratio of coding sequence to 3′ UTR length (Brogna and Wen 2009); the presence of more than one open reading frame in the transcript, including small upstream open reading frames (He *et al.* 2003); the presence of introns in the 3′ UTR (Le Hir *et al.* 2000); and the use of stop codon read-through as a regulatory mechanism (Jungreis *et al.* 2011). Therefore, we asked

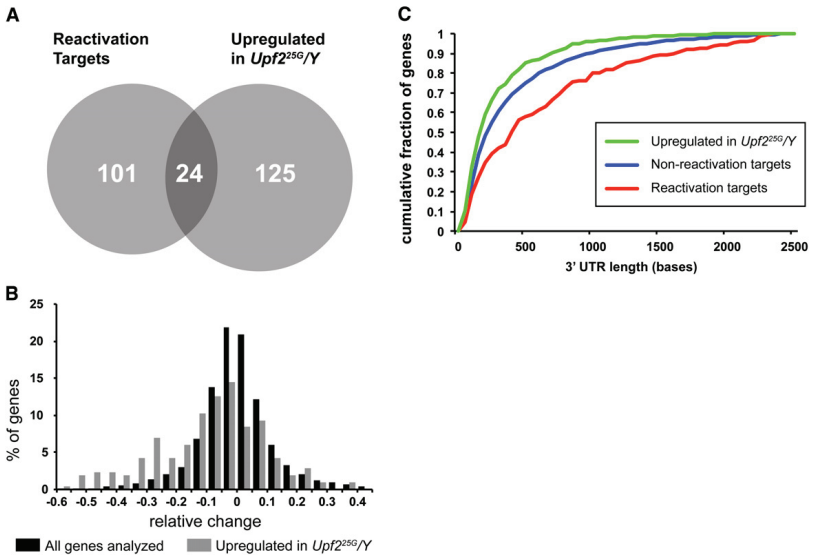


Figure 3 Analysis of reactivation targets. (A) Venn diagram showing overlap of genes that display increased expression in *Upf2*^{25G}*/Y* and reactivation targets. (B) Distribution of relative change in slope (defined in Figure 2) identified in reactivation experiment for all genes (black bars) and genes up-regulated in *Upf2*^{25G}*/Y* (gray bars). (C) Cumulative percentage of 3′ UTR lengths measured in nonreactivation targets genes (5291 genes; blue), reactivation targets (168 genes; red), and genes up-regulated in *Upf2*^{25G}*/Y* (214 genes; green).

whether any of these features are enriched in our NMD reactivation targets compared with nontargets (Table 2). Most strikingly, we found the median length of the 3′ UTR in reactivation targets was 423 bases, considerably longer than the 218 bases found in the nonreactivation control set ($P = 1.55 \times 10^{-5}$ from regression analysis) (Figure 3C). We did not detect significant enrichment for presence of introns in the 3′ UTR, the ratio of length of the coding sequence to 3′ UTR, predicted free energy of the 3′ UTR, read-through candidates, stop codon density in the 3′ UTR, or polycistronic transcripts (Table 2). In contrast to reactivation targets, the 149 mRNAs that show increased expression in the *Upf2*^{25G} mutant have significantly shorter 3′ UTRs on average than genes in our control set (median of 159 bases, $P = 8.71 \times 10^{-8}$; Table 2 and Figure 3C).

Comparison of *in vivo* NMD target analysis to cell-culture analysis

Previously, the most expansive analysis of NMD targeting in *Drosophila* was performed in S2 cell culture (Rehwinkel *et al.* 2005). This analysis identified a set of 184 "core" NMD targets, defined as the up-regulated genes after independent RNAi-mediated silencing of all six known NMD factors in *Drosophila*. Of these 184 genes, 68 passed our data quality control tests, thus allowing us to compare these genes to our NMD-regulated candidates. We find that these overlap between the data sets is relatively small, with the majority genes up-regulated in S2 cells not identified as up-regulated in our *Upf2*^{25G} mutant, or identified as reactivation targets (Figure S3). This discordance between our NMD targeting data collected in whole animals *vs.* data from cell culture experiments is likely a reflection of the tissue, physiological, and developmental stage differences between these two experimental setups.

NMD does not directly regulate genomic load *in vivo*

A proposed role of the NMD pathway is to rid the cell of aberrant transcripts, represented in part by transposable elements (TEs), and sporadic transcripts containing nonsense mutations that arise from errors in transcription or splicing (Mendell *et al.* 2004). A striking finding of our RNA-seq analysis of *Upf2* mutant larvae was a massive increase in reads mapping to the endogenous TE *Copia*. In control *Upf2*$^{+}$ larvae, sequences derived from this LTR-family transposon comprised 0.042% of all mappable reads. However, in *Upf2*^{25G}*/Y*, *Copia* represented 0.77% of all reads (Table S1), a 17.4-fold increase. To test whether *Copia* RNA is a direct target of NMD-mediated degradation, we examined our reactivation samples. Because *Copia* is not represented on the microarrays, we used for global reactivation analysis, we used qRT-PCR to measure *Copia* mRNA in the reactivation time course (Figure 4). *Copia* undergoes an alternative splicing event to generate two major mRNA species (Brierley and Flavell 1990) and we used an qRT-PCR amplicon that detects both of these RNAs (Figure 4A). We found that *Copia* did not behave like a direct NMD pathway target in this assay, as *Copia* levels did not decrease significantly during our reactivation time course (Figure 4B). This conclusion was corroborated from results obtained in S2 cells, where RNAi-mediated depletion of *Upf1* increased *Copia* expression levels but did not alter the decay rate of *Copia* RNA (Figure 4C). Thus, we conclude that *Copia* transcript levels are regulated by processes downstream of direct NMD targeting, as we have

Table 1 The 24 genes up-regulated in *Upf2*^{25G} and identified as reactivation targets, sorted by increased expression observed in mutant

FBgn	Gene Name	Fold Up-Regulation in *Upf2*^{25G}	*P* Value	Relative Slope	Relative Slope *P*-Value	Fold Decay-4 hr
FBgn0040837	*CG8620*	11.84	2.38E-13	−0.26	0.039	2.44
FBgn0034605	*CG15661*	8.84	1.39E-03	−0.43	0.082	2.20
FBgn0033240	*CG2906*	7.30	1.95E-07	−0.08	0.043	1.80
FBgn0029766	*CG15784*	6.52	1.29E-16	−0.40	0.040	2.42
FBgn0261113	*Xrp1*	6.52	3.51E-17	−0.39	0.041	1.94
FBgn0019890	*Smg5*	5.33	5.90E-05	−0.54	0.044	2.44
FBgn0033153	*Gadd45*	5.00	2.23E-06	−0.70	0.040	4.82
FBgn0039319	*CG13659*	4.85	1.84E-10	0.00	0.081	2.37
FBgn0037936	*CG6908*	4.50	1.18E-11	−0.68	0.039	5.21
FBgn0034501	*CG13868*	4.33	9.15E-11	−0.34	0.042	2.12
FBgn0014031	*Spat*	4.19	2.24E-10	−0.15	0.087	2.68
FBgn0031643	*CG3008*	3.94	4.97E-09	−0.14	0.043	1.85
FBgn0041627	*Ku80*	3.58	1.49E-05	−0.26	0.040	1.85
FBgn0032981	*CG3635*	3.34	4.00E-03	−0.19	0.077	2.06
FBgn0037391	*CG2017*	3.34	1.11E-04	−0.20	0.034	2.32
FBgn0050424	*CG30424*	3.31	1.11E-03	−0.46	0.082	2.34
FBgn0039328	*CHKov2*	3.14	3.11E-05	−0.44	0.036	2.66
FBgn0039260	*Smg6*	2.88	3.13E-05	−0.46	0.039	2.07
FBgn0035476	*CG12766*	2.70	1.47E-04	−0.54	0.076	2.16
FBgn0042105	*CG18748*	2.60	3.14E-03	−0.72	0.075	7.18
FBgn0086370	*sra*	2.59	8.02E-05	−0.11	0.082	1.87
FBgn0037781	*Fancl*	2.51	1.58E-03	−0.23	0.083	2.31
FBgn0085194	*CG34165*	2.13	3.54E-03	−0.11	0.049	1.81
FBgn0039226	*Ude*	1.91	7.12E-03	−0.50	0.088	2.20

Fold up-regulation and associated *P* value are obtained from RNA-seq data. "Relative slope" refers to the decay of transcript in *Upf2*-reactivated larvae calculated as a linear regression of $\log_2$ transformed expression level relative to control larvae; see Figure 2. "Fold decay-4 hr" refers to the fold change in transcript level between the preheat shock and 4-hr time point in animals carrying the *Upf2* cDNA. FBgn, FlyBase gene number.

found for the majority of genes that are expressed at increased levels in NMD mutant backgrounds.

To test whether *Upf2*^{25G}/*Y* animals have an overall increase in mRNAs that contain PTCs, we analyzed the proportion of reads in our RNA-seq data that encode for nonsense transcripts (Table S3). We found that such reads comprise 0.048% of all reads mapping to mRNA in *Upf2*^{25G}, compared to 0.046% in controls. Thus, we found no significant difference in the proportion of nonsense reads between *Upf2*^{25G} and control, suggesting that such mutant transcripts do not accumulate in NMD mutants.

DISCUSSION

Although NMD has been established as an important pathway for posttranscriptional gene regulation in eukaryotes, the identities of the

Table 2 Statistical comparison of NMD-sensitive genes compared with controls

	Nonresponding Genes[a]	Up-Regulated in *Upf2*^{25Gc}		Reactivation Targets[e]		Up-Regulated Reactivation Targets[f]	
	Count or Average[b]	Count or Average	*P* Value[d]	Count or Average	*P* Value	Count or Average	*P* Value
Total	5289	215	n/a	154	n/a	24	
3′ UTR length, bases	450.2	275.4	8.7E-08	739.6	1.6E-05	261.4	2.8E-01
3′ UTR structure, kcal/mol/base	−0.024	−0.023	0.1006	−0.028	0.772	−0.028	0.161
CDS/3′ UTR length ratio	0.416	0.389	0.200	0.661	0.601	0.301	0.990
Read-through candidates	113	1	0.358	5	0.700	0	0.520
Bicistronic genes	23	2	6.5E-02	0	0.570	0	1.000
3′ UTR intron present	256	8	0.761	14	4.7E-01	0	0.359
3′ UTR stop codon density	0.067	0.070	0.052	0.066	0.976	0.066	0.859
Polycistronic genes	31	1	0.105	0	0.790	0	0.836
Up-regulated in *Upf2*^{25G}	n/a	n/a	n/a	24	< 2e-16	n/a	n/a

NMD, nonsense-mediated mRNA decay; UTR, untranslated region; CDS, coding DNA sequence; n/a, not applicable.
[a] Genes that neither are up-regulated in *Upf2*^{25G}/*Y* mutants nor undergo increased decay upon *Upf2* reactivation.
[b] Count or average for the indicated feature.
[c] Genes up-regulated in *Upf2*^{25G}/*Y* mutants.
[d] *P* value based on logistic regression compared with nonresponding genes. Compared features are based on FlyBase annotations (R5.45); 3′ UTR length is the length of the longest annotated 3′ UTR in nucleotides; read-through candidates as defined in (Jungreis *et al.* 2011); free energy as calculated using CentroidFold (Hamada *et al.* 2009)
[e] Genes that undergo decay upon *Upf2* reactivation.
[f] Genes both up-regulated in *Upf2*^{25G}/*Y* mutants and that undergo decay upon *Upf2* reactivation. Only genes that were detected in reactivation microarray experiments are tabulated.

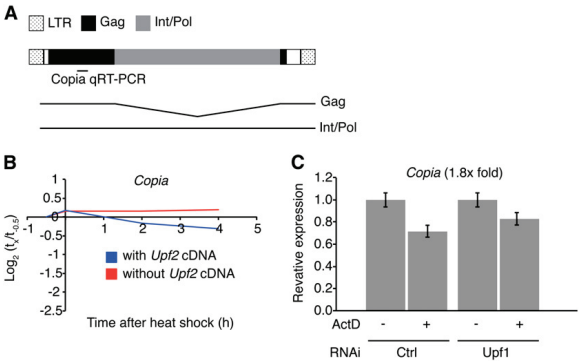


Figure 4 *Copia* RNA levels are indirectly regulated by the NMD pathway. (A) *Copia* genomic (boxes) and transcript structures. Also indicated is the quantitative reverse-transcription polymerase chain reaction (qRT-PCR) amplicon. (B) qRT-PCR analysis of *Copia* of the same time course described in Figure 2. (C) *Copia* levels in S2 cells, measured by qRT-PCR, using the assay described Figure S2 (fold change after *Upf1* depletion indicated in parentheses). Error bars represent ± 1 SD.

directly regulated genes as well as the mechanism of targeting have remained elusive. Identifying the direct targets has been particularly challenging in intact, multicellular organisms. In this work, we have used reactivation of the NMD pathway in an NMD-defective *Drosophila* mutant to identify directly targeted genes. Transcripts that rapidly decrease in abundance upon pathway reactivation are strong candidates for direct targets of NMD-mediated degradation. We have identified 168 candidate direct NMD targets, indicating that NMD regulates a significant proportion of the *Drosophila* genome. This number may, however, be an underestimate due to intrinsic limitations to our approach. For instance, we have examined NMD targeting at one specific stage of development, so we have not identified NMD targeting of genes that are either not transcribed at this stage, or transcribed but not translated, since NMD targeting is translation dependent. Thus, this group of 168 genes likely represents a lower limit to the number of genes directly targeted by the NMD pathway.

In addition to the reactivation assay, we have also measured the change in steady-state expression level of all transcripts in NMD-defective animals. A comparison of the reactivation and steady-state analysis leads to two significant conclusions. First, we found that only 16% of genes that are overexpressed at steady-state levels in the NMD mutant behave like direct targets in our reactivation assay. This finding suggests that, for most transcripts, the increased expression observed in the NMD mutant is due to indirect effects caused by NMD pathway disruption. Such a result was not unexpected, as it is likely that changes in direct target expression will lead to changes in downstream genes, either through specific regulation by overexpressed direct targets, or through changes in organismal development or physiology. In yeast, 45–54% of genes up-regulated in NMD mutants are direct targets (Guan *et al.* 2006; Johansson *et al.* 2007). The finding that in whole *Drosophila*, loss of NMD shows a greater percentage of up-regulated indirect targets may be due to the diversity of tissue-specific physiologic response brought about by loss of NMD (Weischenfeldt *et al.* 2008; Bruno *et al.* 2011; Colak *et al.* 2013). This scenario is important to consider when developing systemically administered therapies designed to modulate NMD activity (Durand *et al.* 2007).

Our second conclusion is more surprising. We found that 81% of candidate direct targets do not show altered steady-state expression in our NMD mutant. This implies there exist feedback mechanisms that act to renormalize the expression of many NMD-target genes when NMD function is compromised. Many gene expression networks contain negative regulatory loops that act to restore expression to physiological levels upon perturbation of pathway components (Becskei and Serrano 2000). This feedback could function indirectly, for instance at the transcriptional level, but in the case of NMD, the feedback could also be working at the level of NMD pathway activity. The NMD effectors *Smg5* and *Smg6* are direct NMD targets, both in mammalian and fly cells (Rehwinkel *et al.* 2005; Huang *et al.* 2011; Yepiskoposyan *et al.* 2011), and our reactivation experiments demonstrate direct targeting of these genes in intact *Drosophila*. The up-regulation of *Smg5* and/or *Smg6* may compensate for the partial loss of *Upf2* function in our *Upf2* hypomorph, restoring gene expression of many direct targets to homeostatic levels. However, upon reactivation of *Upf2*, these direct targets are poised for rapid degradation due to increased expression of *Smg5*, *Smg6*, and other potential NMD factors. Such a model may account for why in yeast a greater percentage of direct NMD target genes show increased expression in NMD mutants (Johansson *et al.* 2007), since there is no evidence for NMD pathway autoregulation in this organism. That some directly targeted genes do show up-regulation in $Upf2^{25G}$ may be indicative of a greater dependence of these genes on core NMD factors for their degradation, as opposed to auxiliary regulators, as has been observed in a number of studies (Chen *et al.* 2005; Metzstein and Krasnow 2006; Avery *et al.* 2011; Huang *et al.* 2011; Frizzell *et al.* 2012; Metze *et al.* 2013). A final speculative possibility is that NMD factors may also be required for transcription of some direct NMD-mediated mRNA decay substrates, as has been shown for yeast proteins involved in general 5′-3′ RNA degradation (Haimovich *et al.* 2013; Sun *et al.* 2013). In this scenario, loss of NMD factors results in reduced transcription of these NMD targets, compensating for the concurrent increase in mRNA stability, which results in unchanged steady-state expression levels. A number of NMD factors are known to shuttle between the cytoplasm and the nucleus (Isken and Maquat 2008), and some are known to have nuclear functions (Brogna 1999; de Turris *et al.* 2011; Varsally and Brogna 2012).

Although our conclusions that most genes up-regulated in NMD-defective *Drosophila* are not direct targets of the NMD pathway, and that many direct targets are not up-regulated, differs what has been found in yeast (Guan *et al.* 2006; Johansson *et al.* 2007), they are similar to findings in mammalian cell culture (Tani *et al.* 2012a). Tani *et al.* (2012a) found that 23% of genes in HeLa cells that are up-regulated under conditions of NMD deficiency appear to be direct NMD targets, similar to the 16% we have shown here. They also

observed that the great majority of genes (90%) that displayed NMD pathway-dependent decay, and thus appeared to be direct degradation targets, had no change in steady-state expression levels when the NMD pathway was inactivated. Again, this is very similar to what we have found in intact *Drosophila*, in which 81% of putative direct targets have no significant change in expression. That such similar results were found using different methodologies in diverse organisms suggests that the phenomenon of feedback regulation normalizing levels of direct NMD targets is a general property of the NMD pathway in metazoans.

Although a rapid decrease in mRNA expression levels upon Upf2 reactivation suggests that the gene is a direct NMD target, it remains possible that the expression levels of some of our identified genes decrease due to a non NMD mechanism. This alternative mechanism is likely to depend on activation of a secondary RNA decay pathway, as transcriptional changes in response to Upf2 expression is unlikely in itself to lead to the rapid transcript decay we observe. Future experiments, such as detection of specific NMD-dependent decay intermediates, could be used to address this alternative possibility; such an approach has been well established in cell culture, but not previously in intact animals.

Most studies seeking to globally identify NMD targets have used overexpression as the main criteria for being an NMD target (Lelivelt and Culbertson 1999; He *et al.* 2003; Mendell *et al.* 2004; Rehwinkel *et al.* 2005; Guan *et al.* 2006; Metzstein and Krasnow 2006; Wittmann *et al.* 2006; Barberan-Soler *et al.* 2009; Ramani *et al.* 2009; McIlwain *et al.* 2010; Nicholson *et al.* 2010; Yepiskoposyan *et al.* 2011; Hurt *et al.* 2013; Matia-González *et al.* 2013). Cataloging NMD targets in this way produces a list with a mix of indirect and direct NMD targets. This in turn, convolutes the correlation between 3′ UTR length and the likelihood of being an *in vivo* NMD target, as has been observed in previous analyses (Rehwinkel *et al.* 2005; Hansen *et al.* 2009; Ramani *et al.* 2009; Yepiskoposyan *et al.* 2011; Rayson *et al.* 2012). In experimental situations, long 3′ UTRs generally make transcripts sensitive to NMD (Kertész *et al.* 2006; Behm-Ansmant *et al.* 2007a; Longman *et al.* 2007; Singh *et al.* 2008; Eberle *et al.* 2009; Hogg and Goff 2010). Our examination of natural *in vivo* NMD targets supports this conclusion: we find that candidate direct NMD targets have considerably longer 3′ UTRs on average than nondirect targets. Almost 50% of candidate direct target transcripts have a 3′ UTR over 420 bases in length, a value suggested to be a threshold between NMD resistance and sensitivity (Singh *et al.* 2008), compared with only 31% of nontargeted transcripts. However, as has been found in mammalian cells, we find that a large number of transcripts that have long 3′ UTRs are not subject to NMD (with the caveat that, in most cases, defining a transcript as NMD sensitive in mammalian cells is based on increased expression when the NMD pathway is disabled). Singh *et al.* (2008) have proposed such long but NMD-insensitive transcripts contain specific NMD-protective sequences. Conversely, we also find that there are many genes with short 3′ UTRs that appear to be direct NMD targets. Since most models invoke a long 3′ UTR as the trigger for NMD targeting, how structurally normal, short 3′ UTRs could be targeted is unclear. We propose these transcripts may contain specific sequence features that stimulate NMD targeting, analogous to yeast downstream elements (Peltz *et al.* 1993; González *et al.* 2001). Although such elements are not required to trigger decay in NMD-sensitive transcripts with long 3′ UTRs, they could be key to targeting mRNAs with short 3′ UTRs. Furthermore, the decay rates of both mutant PTC-containing and endogenous transcripts are often quite variable when measured experimentally, a phenomenon that could be explained by the presence or absence of DSE-like elements in these transcripts.

Finally, as described previously, our results regarding the number of direct targets of the NMD pathway correlate well with the findings in HeLa cells. However, the results from this study differed from ours in one significant aspect: in HeLa cells direct targets were found to have shorter 3′ UTR lengths than nontargets (Tani *et al.* 2012a), a finding contradictory to our own and unexpected given current models of NMD targeting. A possible explanation for this discrepancy is that cell transformation is associated with a global alteration in 3′ UTR length (Mayr and Bartel 2009). Potentially, the 3′ UTR lengths of NMD targeted transcripts in HeLa cells systematically differ from the annotated human genome reference set, explaining the discordance. It will be important in the future to examine NMD targeting in nontransformed mammalian cells to determine the contribution of 3′ UTR length to NMD sensitivity.

One model for why the NMD pathway is required for viability is that NMD functions to suppress transcriptome noise, such as the noise arising from endogenous TEs or erroneously generated PTC-bearing mRNAs. However, although we find that the TE *Copia* is highly up-regulated in $Upf2^{25G}$ mutants, Copia RNA is not a direct NMD target *in vivo*. Therefore, *Copia* up-regulation is likely due to processes downstream of NMD, such as stress responses, which are known to up-regulate *Copia* in other experimental situations (Strand and McDonald 1985). Moreover, we find no significant enrichment in reads that map to PTC-generating transcripts in $Upf2^{25G}$, suggesting that NMD does not function to suppress such transcripts *in vivo*. Since we find no evidence that NMD suppresses genomic noise *in vivo*, we favor a model in which NMD is required for viability through regulation of particular critical NMD target genes. Therefore, the understanding of the *in vivo* roles of the NMD pathway will require the identification of these targets, particularly those that display increased expression in NMD mutants, as it is likely that this increased expression mediates the biological defects observed in NMD-defective organisms and cells. Our analysis indicates that this is a relatively short list; we find only 24 genes appear to be both be overexpressed and be direct targets at the developmental stage we examined. The continued analysis of these 24 genes should reveal important insights into NMD-dependent roles in normal development and physiology.

ACKNOWLEDGMENTS

We thank the Metzstein, Yandell, and Thummel labs for helpful discussions; Barry Moore, Zev Kronenberg, and E. J. Osborne for bioinformatics support; Deepika Gaddam for S2 cell culture; the GNomEX bioinformatics core at the University of Utah Huntsman Cancer Institute; and Carl Thummel, Nels Elde, Kim Frizzell, and members of the Metzstein laboratory for comments on the manuscript. Fly stocks were obtained from the Bloomington *Drosophila* stock center. This work was supported by National Institutes of Health grant 1R01GM084011 (to M.M.M. and M.Y.) and a March of Dimes Award 5-FY07-664 (to M.M.M.).

LITERATURE CITED

Anders, S., and W. Huber, 2010 Differential expression analysis for sequence count data. Genome Biol. 11: R106.

Avery, P., M. Vicente-Crespo, D. Francis, O. Nashchekina, C. R. Alonso *et al.*, 2011 *Drosophila* Upf1 and Upf2 loss of function inhibits cell growth and causes animal death in a Upf3-independent manner. RNA 17: 624–638.

Azzalin, C. M., and J. Lingner, 2006 The human RNA surveillance factor UPF1 is required for S phase progression and genome stability. Curr. Biol. 16: 433–439.

Barberan-Soler, S., N. J. Lambert, and A. M. Zahler, 2009 Global analysis of alternative splicing uncovers developmental regulation of nonsense-mediated decay in *C. elegans*. RNA 15: 1652–1660.

Becskei, A., and L. Serrano, 2000 Engineering stability in gene networks by autoregulation. Nature 405: 590–593.

Behm-Ansmant, I., D. Gatfield, J. Rehwinkel, V. Hilgers, and E. Izaurralde, 2007a A conserved role for cytoplasmic poly(A)-binding protein 1 (PABPC1) in nonsense-mediated mRNA decay. EMBO J. 26: 1591–1601.

Behm-Ansmant, I., I. Kashima, J. Rehwinkel, J. Saulière, N. Wittkopp *et al.*, 2007b mRNA quality control: an ancient machinery recognizes and degrades mRNAs with nonsense codons. FEBS Lett. 581: 2845–2853.

Brand, A. H., and N. Perrimon, 1993 Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development 118: 401–415.

Brierley, C., and A. J. Flavell, 1990 The retrotransposon copia controls the relative levels of its gene products post-transcriptionally by differential expression from its two major mRNAs. Nucleic Acids Res. 18: 2947–2951.

Brogna, S., 1999 Nonsense mutations in the alcohol dehydrogenase gene of *Drosophila* melanogaster correlate with an abnormal 3′ end processing of the corresponding pre-mRNA. RNA 5: 562–573.

Brogna, S., and J. Wen, 2009 Nonsense-mediated mRNA decay (NMD) mechanisms. Nat. Struct. Mol. Biol. 16: 107–113.

Bruno, I. G., R. Karam, L. Huang, A. Bhardwaj, C. H. Lou *et al.*, 2011 Identification of a microRNA that activates gene expression by repressing nonsense-mediated RNA decay. Mol. Cell 42: 500–510.

Chen, Z., K. R. Smith, P. Batterham, and C. Robin, 2005 Smg1 nonsense mutations do not abolish nonsense-mediated mRNA decay in *Drosophila melanogaster*. Genetics 171: 403–406.

Chomczynski, P., and N. Sacchi, 1987 Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162: 156–159.

Colak, D., S.-J. Ji, B. T. Porse, and S. R. Jaffrey, 2013 Regulation of axon guidance by compartmentalized nonsense-mediated mRNA decay. Cell 153: 1252–1265.

de Turris, V., P. Nicholson, R. Z. Orozco, R. H. Singer, and O. Muhlemann, 2011 Cotranscriptional effect of a premature termination codon revealed by live-cell imaging. RNA 17: 2094–2107.

Durand, S., N. Cougot, F. Mahuteau-Betzer, C.-H. Nguyen, D. S. Grierson *et al.*, 2007 Inhibition of nonsense-mediated mRNA decay (NMD) by a new chemical molecule reveals the dynamic of NMD factors in P-bodies. J. Cell Biol. 178: 1145–1160.

Eberle, A. B., L. Stalder, H. Mathys, R. Z. Orozco, and O. Mühlemann, 2008 Posttranscriptional gene regulation by spatial rearrangement of the 3′ untranslated region. PLoS Biol. 6: e92.

Eberle, A. B., S. Lykke-Andersen, O. Mühlemann, and T. H. Jensen, 2009 SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. Nat. Struct. Mol. Biol. 16: 49–55.

Frizzell, K. A., S. G. Rynearson, and M. M. Metzstein, 2012 *Drosophila* mutants show NMD pathway activity is reduced, but not eliminated, in the absence of Smg6. RNA 18: 1475–1486.

González, C. I., A. Bhattacharya, W. Wang, and S. W. Peltz, 2001 Nonsense-mediated mRNA decay in *Saccharomyces cerevisiae*. Gene 274: 15–25.

Guan, Q., W. Zheng, S. Tang, X. Liu, R. A. Zinkel *et al.*, 2006 Impact of nonsense-mediated mRNA decay on the global expression profile of budding yeast. PLoS Genet. 2: e203.

Haimovich, G., D. A. Medina, S. Z. Causse, M. Garber, G. Millán-Zambrano *et al.*, 2013 Gene expression is circular: factors for mRNA degradation also foster mRNA synthesis. Cell 153: 1000–1011.

Hamada, M., H. Kiryu, K. Sato, T. Mituyama, and K. Asai, 2009 Prediction of RNA secondary structure using generalized centroid estimators. Bioinformatics 25: 465–473.

Hansen, K. D., L. F. Lareau, M. Blanchette, R. E. Green, Q. Meng *et al.*, 2009 Genome-wide identification of alternative splice forms down-regulated by nonsense-mediated mRNA decay in *Drosophila*. PLoS Genet. 5: e1000525.

He, F., X. Li, P. Spatrick, R. Casillo, S. Dong *et al.*, 2003 Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5′ to 3′ mRNA decay pathways in yeast. Mol. Cell 12: 1439–1452.

Hodgkin, J., A. Papp, R. Pulak, V. Ambros, and P. Anderson, 1989 A new kind of informational suppression in the nematode *Caenorhabditis elegans*. Genetics 123: 301–313.

Hogg, J. R., and S. P. Goff, 2010 Upf1 senses 3′UTR length to potentiate mRNA decay. Cell 143: 379–389.

Huang, L., C.-H. Lou, W. Chan, E. Y. Shum, A. Shao *et al.*, 2011 RNA homeostasis governed by cell type-specific and branched feedback loops acting on NMD. Mol. Cell 43: 950–961.

Hurt, J. A., A. D. Robertson, and C. B. Burge, 2013 Global analyses of UPF1 binding and function reveals expanded scope of nonsense-mediated mRNA decay. Genome Res. 23: 1636–1650.

Hwang, J., and L. E. Maquat, 2011 Nonsense-mediated mRNA decay (NMD) in animal embryogenesis: to die or not to die, that is the question. Curr. Opin. Genet. Dev. 21: 422–430.

Isken, O., and L. E. Maquat, 2008 The multiple lives of NMD factors: balancing roles in gene and genome regulation. Nat. Rev. Genet. 9: 699–712.

Johansson, M. J. O., F. He, P. Spatrick, C. Li, and A. Jacobson, 2007 Association of yeast Upf1p with direct substrates of the NMD pathway. Proc. Natl. Acad. Sci. USA 104: 20872–20877.

Jungreis, I., M. F. Lin, R. Spokony, C. S. Chan, N. Negre *et al.*, 2011 Evidence of abundant stop codon readthrough in *Drosophila* and other Metazoa. Genome Res. 21: 2096–2113.

Kebaara, B. W., and A. L. Atkin, 2009 Long 3′-UTRs target wild-type mRNAs for nonsense-mediated mRNA decay in *Saccharomyces cerevisiae*. Nucleic Acids Res. 37: 2771–2778.

Kerényi, Z., Z. Mérai, L. Hiripi, A. Benkovics, P. Gyula *et al.*, 2008 Inter-kingdom conservation of mechanism of nonsense-mediated mRNA decay. EMBO J. 27: 1585–1595.

Kertész, S., Z. Kerényi, Z. Mérai, I. Bartos, T. Pálfy *et al.*, 2006 Both introns and long 3′-UTRs operate as cis-acting elements to trigger nonsense-mediated decay in plants. Nucleic Acids Res. 34: 6147–6157.

Kervestin, S., and A. Jacobson, 2012 NMD: a multifaceted response to premature translational termination. Nat. Rev. Mol. Cell Biol. 13: 700–712.

Le Hir, H., E. Izaurralde, L. E. Maquat, and M. J. Moore, 2000 The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon-exon junctions. EMBO J. 19: 6860–6869.

Leeds, P., S. W. Peltz, A. Jacobson, and M. R. Culbertson, 1991 The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon. Genes Dev. 5: 2303–2314.

Lelivelt, M. J., and M. R. Culbertson, 1999 Yeast Upf proteins required for RNA surveillance affect global expression of the yeast transcriptome. Mol. Cell. Biol. 19: 6710–6719.

Lloyd, J. P. B., and B. Davies, 2013 SMG1 is an ancient nonsense-mediated mRNA decay effector. Plant J. 76: 800–810.

Longman, D., R. Plasterk, I. Johnstone, and J. Caceres, 2007 Mechanistic insights and identification of two novel factors in the *C. elegans* NMD pathway. Genes Dev. 21: 1075–1085.

Maderazo, A. B., J. P. Belk, F. He, and A. Jacobson, 2003 Nonsense-containing mRNAs that accumulate in the absence of a functional nonsense-mediated mRNA decay pathway are destabilized rapidly upon its restitution. Mol. Cell. Biol. 23: 842–851.

Matia-González, A. M., A. Hasan, G. H. Moe, J. Mata, and M. A. Rodríguez-Gabriel, 2013 Functional characterization of Upf1 targets in *Schizosaccharomyces pombe*. RNA Biol. 10: 1057–1065.

Mayr, C., and D. P. Bartel, 2009 Widespread shortening of 3′UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. Cell 138: 673–684.

McIlwain, D. R., Q. Pan, P. T. Reilly, A. J. Elia, S. McCracken *et al.*, 2010 Smg1 is required for embryogenesis and regulates diverse genes via alternative splicing coupled to nonsense-mediated mRNA decay. Proc. Natl. Acad. Sci. USA 107: 12186–12191.

Medghalchi, S. M., P. A. Frischmeyer, J. T. Mendell, A. G. Kelly, A. M. Lawler *et al.*, 2001 Rent1, a trans-effector of nonsense-mediated mRNA decay,

is essential for mammalian embryonic viability. Hum. Mol. Genet. 10: 99–105.
Mendell, J. T., S. M. Medghalchi, R. G. Lake, E. N. Noensie, and H. C. Dietz, 2000 Novel Upf2p orthologues suggest a functional link between translation initiation and nonsense surveillance complexes. Mol. Cell. Biol. 20: 8944–8957.
Mendell, J. T., N. A. Sharifi, J. L. Meyers, F. Martinez-Murillo, and H. C. Dietz, 2004 Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nat. Genet. 36: 1073–1078.
Metze, S., V. A. Herzog, M.-D. Ruepp, and O. Mühlemann, 2013 Comparison of EJC-enhanced and EJC-independent NMD in human cells reveals two partially redundant degradation pathways. RNA 19: 1432–1448.
Metzstein, M. M., and M. A. Krasnow, 2006 Functions of the nonsense-mediated mRNA decay pathway in *Drosophila* development. PLoS Genet. 2: e180.
Neu-Yilik, G., N. Gehring, M. Hentze, and A. Kulozik, 2004 Nonsense-mediated mRNA decay: from vacuum cleaner to Swiss army knife. Genome Biol. 5: 218.
Nicholson, P., H. Yepiskoposyan, S. Metze, R. Zamudio Orozco, N. Kleinschmidt *et al.*, 2010 Nonsense-mediated mRNA decay in human cells: mechanistic insights, functions beyond quality control and the double-life of NMD factors. Cell. Mol. Life Sci. 67: 677–700.
Nix, D. A., S. J. Courdy, and K. M. Boucher, 2008 Empirical methods for controlling false positives and estimating confidence in ChIP-Seq peaks. BMC Genomics 9: 523.
Palacios, I. M., 2013 Nonsense-mediated mRNA decay: from mechanistic insights to impacts on human health. Brief. Funct. Genomics 12: 25–36.
Peltz, S. W., A. H. Brown, and A. Jacobson, 1993 mRNA destabilization triggered by premature translational termination depends on at least three cis-acting sequence elements and one trans-acting factor. Genes Dev. 7: 1737–1754.
Ramani, A. K., A. C. Nelson, P. Kapranov, I. Bell, T. R. Gingeras *et al.*, 2009 High resolution transcriptome maps for wild-type and nonsense-mediated decay-defective *Caenorhabditis elegans*. Genome Biol. 10: R101.
Rayson, S., L. Arciga-Reyes, L. Wootton, M. De Torres Zabala, W. Truman *et al.*, 2012 A role for nonsense-mediated mRNA decay in plants: pathogen responses are induced in *Arabidopsis thaliana* NMD mutants. PLoS ONE 7: e31917.
Rehwinkel, J., I. Letunic, J. Raes, P. Bork, and E. Izaurralde, 2005 Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. RNA 11: 1530–1544.
Riehs, N., S. Akimcheva, J. Puizina, P. Bulankova, R. A. Idol *et al.*, 2008 Arabidopsis SMG7 protein is required for exit from meiosis. J. Cell Sci. 121: 2208–2216.
Roberts, T. L., U. Ho, J. Luff, C. S. Lee, S. H. Apte *et al.*, 2013 Smg1 haploinsufficiency predisposes to tumor formation and inflammation. Proc. Natl. Acad. Sci. USA 110: E285–E294.
Rodríguez-Gabriel, M. A., S. Watt, J. Bähler, and P. Russell, 2006 Upf1, an RNA helicase required for nonsense-mediated mRNA decay, modulates the transcriptional response to oxidative stress in fission yeast. Mol. Cell. Biol. 26: 6347–6356.
Roman, G., J. He, and R. L. Davis, 2000 kurtz, a novel nonvisual arrestin, is an essential neural gene in *Drosophila*. Genetics 155: 1281–1295.
Schweingruber, C., S. C. Rufener, D. Zünd, A. Yamashita, and O. Mühlemann, 2013 Nonsense-mediated mRNA decay - Mechanisms of substrate mRNA recognition and degradation in mammalian cells. Biochim. Biophys. Acta. 1829: 612–623.
Singh, G., I. Rebbapragada, and J. Lykke-Andersen, 2008 A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay. PLoS Biol. 6: e111.
Smyth, G. K., and T. Speed, 2003 Normalization of cDNA microarray data. Methods 31: 265–273.
Stapleton, M., G. Liao, P. Brokstein, L. Hong, P. Carninci *et al.*, 2002 The *Drosophila* gene collection: identification of putative full-length cDNAs for 70% of *D. melanogaster* genes. Genome Res. 12: 1294–1300.
Strand, D. J., and J. F. McDonald, 1985 Copia is transcriptionally responsive to environmental stress. Nucleic Acids Res. 13: 4401–4410.
Sullivan, W., M. Ashburner, and R. S. Hawley, 2000 *Drosophila Protocols*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sun, M., B. Schwalb, N. Pirkl, K. C. Maier, A. Schenk *et al.*, 2013 Global analysis of eukaryotic mRNA degradation reveals xrn1-dependent buffering of transcript levels. Mol. Cell 52: 52–62.
Tani, H., N. Imamachi, K. A. Salam, R. Mizutani, K. Ijiri *et al.*, 2012a Identification of hundreds of novel UPF1 target transcripts by direct determination of whole transcriptome stability. RNA Biol. 9: 1370–1379.
Tani, H., R. Mizutani, K. A. Salam, K. Tano, K. Ijiri *et al.*, 2012b Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals. Genome Res. 22: 947–956.
Trapnell, C., A. Roberts, L. Goff, G. Pertea, D. Kim *et al.*, 2012 Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. 7: 562–578.
Varsally, W., and S. Brogna, 2012 UPF1 involvement in nuclear functions. Biochem. Soc. Trans. 40: 778–783.
Viegas, M. H., N. H. Gehring, S. Breit, M. W. Hentze, and A. E. Kulozik, 2007 The abundance of RNPS1, a protein component of the exon junction complex, can determine the variability in efficiency of the nonsense mediated decay pathway. Nucleic Acids Res. 35: 4542–4551.
Weischenfeldt, J., I. Damgaard, D. Bryder, K. Theilgaard-Mönch, L. A. Thoren *et al.*, 2008 NMD is essential for hematopoietic stem and progenitor cells and for eliminating by-products of programmed DNA rearrangements. Genes Dev. 22: 1381–1396.
Weischenfeldt, J., J. Waage, G. Tian, J. Zhao, I. Damgaard *et al.*, 2012 Mammalian tissues defective in nonsense-mediated mRNA decay display highly aberrant splicing patterns. Genome Biol. 13: R35.
Wittkopp, N., E. Huntzinger, C. Weiler, J. Saulière, S. Schmidt *et al.*, 2009 Nonsense-mediated mRNA decay effectors are essential for zebrafish embryonic development and survival. Mol. Cell. Biol. 29: 3517–3528.
Wittmann, J., E. M. Hol, and H.-M. Jäck, 2006 hUPF2 silencing identifies physiologic substrates of mammalian nonsense-mediated mRNA decay. Mol. Cell. Biol. 26: 1272–1287.
Xu, T., and G. Rubin, 1993 Analysis of genetic mosaics in developing and adult *Drosophila* tissues. Development 117: 1223–1237.
Yepiskoposyan, H., F. Aeschimann, D. Nilsson, M. Okoniewski, and O. Mühlemann, 2011 Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. RNA 17: 2108–2118.
Zünd, D., A. R. Gruber, M. Zavolan, and O. Mühlemann, 2013 Translation-dependent displacement of UPF1 from coding sequences causes its enrichment in 3′ UTRs. Nat. Struct. Mol. Biol. 20: 936–943.

Communicating editor: H. D. Lipshitz

CHAPTER 3

GADD45 IS A CRITICAL DIRECT TARGET OF THE NMD PATHWAY

Abstract

Although NMD is required for viability in most organisms, it is not known how it exerts its physiological roles. In this chapter, we exploit the results of an unbiased screen for effectors of the NMD loss-of-function phenotype to identify *Gadd45* as a physiologically relevant NMD target. *Gadd45* is upregulated in NMD mutants and is a direct NMD target *in vivo*. We have also generated the first mutants for *Gadd45* in *Drosophila* and find that disruption of *Gadd45* can ameliorate aspects of the NMD null mutant phenotype, such as developmental arrest. In conclusion, our results suggest that direct, negative regulation of *Gadd45* by the NMD pathway is critical to organismal development and survival.

Introduction

In most organisms including *Drosophila*, zebrafish, mammals, and plants, NMD pathway genes are required for viability (Medghalchi et al., 2001; Metzstein and Krasnow, 2006; Weischenfeldt et al., 2008; Wittkopp et al., 2009; Yoine et al., 2006). While NMD genes are not necessary for viability in yeasts and *C. elegans*, they are

required for the response to stress and for normal development of the reproductive organs, respectively (Cali et al., 1999; Hodgkin et al., 1989; Leeds et al., 1992; Matia-González et al., 2013). It is likely that function of the NMD pathway itself is required for development and physiology. For instance, within each organism, the loss of any individual NMD gene results in remarkably similar phenotypes, suggesting that it is NMD pathway function itself that has physiologically critical functions (Hwang and Maquat, 2011). In cases where different mutations produce a range of phenotypes, the severity of the NMD defect is typically correlated with the severity of other physiological and developmental defects (Frizzell et al., 2012; Metzstein and Krasnow, 2006). Additionally, NMD is required for wild-type regulation of gene expression due to the ability of the NMD pathway to directly destabilize endogenous target genes (Neu-Yilik et al., 2004). Taken together, these observations suggest that the NMD pathway is an important component of eukaryotic biology, and in NMD mutants, accumulation of native target genes interferes with normal development and physiology.

While it is clear that NMD functions as a posttranscriptional regulator of physiologically relevant biological processes, in the vast majority of cases it is not known which NMD target genes are relevant to these processes (Johansson and Jacobson, 2010; Matia-González et al., 2013; Medghalchi et al., 2001; Metzstein and Krasnow, 2006; Weischenfeldt et al., 2008; Wittkopp et al., 2009; Yoine et al., 2006). Furthermore, there have been no unbiased screens designed to identify suppressors of NMD mutant phenotypes. Also, in most organisms, it is not known which of the genes upregulated in NMD mutants are direct targets and which are upregulated as a secondary consequences to misregulation of directly targeted transcripts. Distinguishing direct from indirect

targets is important to understanding the critical biological roles of NMD, as direct targets of NMD are presumed to control expression of secondary targets.

In *Drosophila, Upf1, Upf2* and *Smg5* are required for viability, with mutant larvae dying in midlarval or pupal stages (Cali et al., 1999; Chapin et al., 2014; Hodgkin et al., 1989; Leeds et al., 1992; Matia-González et al., 2013; Metzstein and Krasnow, 2006)(J.O. Nelson, unpublished data). NMD is also required for cell proliferation with NMD-mutant tissue being defective in cell/cell competition assays (Avery et al., 2011; Hwang and Maquat, 2011; Metzstein and Krasnow, 2006; Rehwinkel et al., 2005). Recently, we have identified a set of genes that are overexpressed in *Drosophila* NMD mutants and are direct targets of the NMD pathway based on a functional reactivation assay (Chapter 2 of this thesis and in Chapin et al., 2014). It is likely that one or more of these genes mediate the defects observed in NMD loss-of-function mutants.

Here, we present evidence that one of these 24 genes, *Gadd45*, is a critical downstream mediator of NMD mutant defects. We find that reducing the activity of *Gadd45*, or its critical downstream effector, *Mekk1*, can ameliorate the loss of NMD pathway activity, including developmental progression and cell proliferation. Our conclusions, supported by data presented in this study using *Drosophila,* may be applicable across taxa as *Gadd45* homologues are also directly targeted by NMD in mammalian lineages (Frizzell et al., 2012; Metzstein and Krasnow, 2006; Viegas et al., 2007).

Results

A deficiency rescue screen for genes that suppress the lethality associated with NMD dysfunction identifies *Gadd45* and *Mekk1*

Our approach took advantage of a screen previously conducted in the lab to identify genes with a functional, epistatic relationship to the NMD pathway (Materials and methods). This screen was based on the ability of heterozygous deficiencies to ameliorate the partial inviability of a strong *Upf2* hypomorph, $Upf2^{25G}$. We tested a collection of 413 isogenic deficiencies for those that, when heterozygous, significantly increase the number of viable $Upf2^{25G}/Y$ males above their 10% survival rate (Figure 3.1 and Chapin et al*.,* 2014). To quantify the ability of individual deficiencies to rescue $Upf2^{25G}/Y$, we developed a metric we call the "Deficiency Rescue Score" (DRS, see Materials and methods). Deficiencies with a DRS > 0 show suppression of $Upf2^{25G}/Y$*,* with scores > 0.1 considered significant. These deficiencies are thus candidates for containing critical NMD target genes (Figure 3.2).

We first examined deficiency coverage of our 24 candidate target genes (Table 3.1). Of these 24 genes, 9 were represented in our deficiency collection (14 total deficiencies). For three genes, *Xrp1*, *CG18784,* and *Gadd45*, all uncovering deficiencies had positive scores. However, for both *Xrp1* and *CG18784,* at least one of these deficiencies had DRS scores below our cutoff, suggesting that the rescuing deficiencies that uncover *Xrp1* or *CG18784* may represent false positives. In contrast, $Upf2^{25G}/Y$ is suppressed effectively by *Df(2R)ED1673,* a deficiency that uncovers the NMD target gene *Gadd45*. Two other deficiencies that do not rescue $Upf2^{25G}/Y$, *Df(2R)ED*1618 and

*Df(2R)ED*1715, flank *Gadd45* but do not disrupt this locus, further supporting the notion that *Gadd45* is the critical target gene uncovered by *Df(2R)ED1673* (Figure 3.3A). This screen also identified two overlapping deficiencies, *Df(3R)ED2* and *Df(3R)ED5911*, that uncover *Mekk1* (Figure 3.3B). Interestingly, the deficiencies that uncover *Mekk1* can also rescue a cell proliferation defect observed in NMD mutants (Figure 3.4 and Materials and Methods).

Gadd45 disruption can ameliorate NMD loss of function defects

To test the hypothesis that *Gadd45* is the rescuing gene uncovered by *Df(2R)ED1673*, we generated mutant alleles of *Gadd45* using TALEN-mediated mutagenesis (Figure 3.5 and Materials and Methods). This procedure generated a mutant allele with a disruption to the *Gadd45* coding sequence, $Gadd45^{54C}$. This allele eliminates six in-frame amino acids, does not disrupt the reading frame, and therefore is not predicted to be a null allele. As a control, we also isolated a chromosome that had gone through the TALEN procedure but has no detectable change in *Gadd45,* $Gadd45^{41B}$. To test if *Gadd45* mediates the NMD-mutant phenotype, we asked if $Gadd45^{54C}$ behaves similarly to *Df(2R)ED1673,* which partially rescues $Upf2^{25G}/Y$ as a heterozygote. We found that $Gadd45^{54C}$ partially rescued viability in $Upf2^{25G}/Y$, while $Gadd45^{41B}$ did not (Figure 3.6A). However, rescue by $Gadd45^{54C}$ is to a lesser degree than *Df(2R)ED1673*, suggesting that $Gadd45^{54C}$ is a hypomorphic allele (as predicted from sequence analysis), or that other important loci exist within the rescuing deficiency.

We next asked if *Gadd45* disruption can rescue the predominantly L2 lethal phase

of NMD null mutant animals (Chapin et al., 2014). We analyzed the lethal phase of *Upf2*^{14J} mutants that were also homozygous for either *Gadd45*^{41B} or *Gadd45*^{54C}. This analysis revealed that ~80% of *Upf2*^{14J}*; Gadd45*^{41B} larvae arrest and die in the L2 stage, with the remainder dying in L1; a similar lethal phase to which we observed for this allele of *Upf2* (Chapin et al., 2014)(Figure 3.6B). This lethal phase is in contrast to *Upf2*^{14J}*/Y; Gadd45*^{54C}, which die predominantly as L3 (Figure 3.6B). The partial rescue of the NMD-mutant lethal phase from L2 to L3 by the disruption of *Gadd45* indicates that *Gadd45* overexpression in an NMD mutant inhibits development past the L2 stage.

Loss of *Mekk1* suppresses NMD-mediated defects

To identify the role of *Mekk1* in the NMD mutant phenotype, we tested a loss of function allele known as *Ur36* (Inoue et al., 2001). This allele was generated using imprecise excision techniques and removes 868 bp coding for the kinase domain and is thought to be a null. We crossed animals harboring *Mekk1*^{Ur36} to *Upf2*^{25G}*/FM7i* animals and assayed for viability of F1 *Upf2*^{25G}*/Y; Mekk1*^{Ur36}/+ animals. We found that *Mekk1*^{Ur36} partially rescues viability, similar to that of the rescuing deficiency that uncovers the *Mekk1* locus (Figure 3.7), and suggests that *Mekk1* is the rescuing gene uncovered by *Df(3R)ED2* and *Df(3R)ED5911.* As *Mekk1* is known to be an important downstream effector of *Gadd45*, our result that *Mekk1* also appears to mediate the NMD loss-of-function phenotype and strongly implicates the *Gadd45*/*Mekk1* signaling pathway as a critical target pathway of NMD.

A large portion of the genes regulated by *Mekk1* are secondary targets of the NMD pathway

Mekk1 is transcriptionally activated in response to epithelial wounding, and the expression of several genes involved in the wounding response are upregulated in a *Mekk1* dependent manner (Brun et al., 2006). We asked if the transcriptional output of *Mekk1* activation resembled gene expression changes brought about by a loss of NMD function. To test this hypothesis, we compared the list of *Mekk1*-dependent wounding genes to those upregulated in $Upf2^{25G}/Y$ animals (Brun et al., 2006; Chapin et al., 2014)(Table 3.2). Of the 13 genes upregulated in a *Mekk1*-dependent manner upon wounding, six are upregulated in > 1.5 fold in $Upf2^{25G}/Y$ mutants. Also, none of these six genes are reactivation targets, suggesting that NMD indirectly negatively regulates a large proportion of genes that are also activated by *Mekk1*.

The 3' UTR of *Gadd45* is targeted by NMD in intact larvae

To test the hypothesis that it is the long 3' UTR of *Gadd45* that renders the transcript sensitive to NMD, we constructed a *UAS/eGFP* reporter using the *Gadd45* 3' UTR. We then asked if this construct is susceptible to NMD-mediated decay (Figure 3.8A). When we used *e22c:GAL4* to drive expression of the transgene in both wild-type and $Upf2^{25G}$ mutant backgrounds, we found GFP expression was much higher in $Upf2^{25G}/Y$ animals compared to $Upf2^{+}/Y$ (Figure 3.8B). It is known that NMD does not affect *GAL4* transcription (Metzstein and Krasnow, 2006), thus the changes in GFP expression we observe are most likely due to differences in transcript stability. This result

suggests that the *Gadd45* 3' UTR is sufficient to induce NMD.

Discussion

The identification of the critical target genes of NMD (i.e., native genes which require regulated degradation by the NMD pathway to allow for normal development) is important to understanding the overall function of the NMD pathway *in vivo*. While many NMD targets have been identified, none have previously been shown to be causative of the NMD mutant phenotype. Indeed, until our work, no assay had yet been performed to distinguish direct from indirect NMD targeting genome-wide in an intact, multicellular organism (Chapin et al., 2014). Additionally, no functional screens have been performed to identify relevant NMD target genes. Using two approaches, one based on the identification of direct targets *in vivo* and one based on the suppression of NMD mutant phenotypes, we have identified *Gadd45* as a candidate direct target that functions in mediating aspects of the NMD mutant phenotype.

Interestingly, there is evidence that regulation of *Gadd45* transcript levels by the NMD pathway is evolutionally conserved. *GADD45β* (also called *MyD118),* one of three *Gadd45* homologues found in mammals (Takekawa and Saito, 1998), has been shown to be a direct NMD target, possibly via an intron present in its 3' UTR (Viegas et al., 2007). *Drosophila Gadd45* does not contain an intron in its 3' UTR and indeed splicing, unlike in mammals, is not thought to be crucial for NMD targeting in insects. However, the 3' UTR of *Drosophila Gadd45* is relatively long (743 bases compared to the genome-wide median of 460 bases) a length we have shown is sufficient to render a reporter construct sensitive to NMD. Long 3' UTRs in *Drosophila* and other organisms are a predisposing

feature that increases the likelihood of targeting by NMD (Behm-Ansmant et al., 2007).

As NMD negatively regulates *Gadd45*, it is predicted that any gain-of-function, or activation, phenotype of *Gadd45* will resemble the NMD loss-of-function phenotype. Phenotypes associated with activation of GADD45-family proteins and the observed defects of NMD mutants suggest that *Gadd45* may be an important target of NMD. GADD (Growth arrest and DNA damage) family genes were identified in mammalian cell culture as mRNAs that are activated following treatment with DNA damaging agents, such as UV light, that lead to cellular growth arrest (Fornace et al., 1988; Papathanasiou et al., 1991). Activation of *Gadd45β*, one of three *Gadd45*-family mRNAs appears to be physiologically relevant as overexpression of *Gadd45* in mammalian cell culture is sufficient to induce cell cycle arrest (Jin et al., 2002; Takekawa and Saito, 1998) and in certain circumstances, apoptosis (Mak and Kültz, 2004). These phenotypes mirror those of NMD inhibition in mammals, where silencing of *Upf1* in cell culture induces cell cycle arrest (Azzalin and Lingner, 2006), and complete loss of *Upf1* function in mouse cells induces apoptosis (McIlwain et al., 2010; Medghalchi et al., 2001).

The parallel phenotypes between *Gadd45* activation and NMD loss of function are also recapitulated in *Drosophila*. siRNA-mediated inhibition of the six major *trans*-acting NMD factors in *Drosophila* S2 cell culture results in cell cycle inhibition (Rehwinkel et al., 2005). Moreover, null mutants for *Upf1, Upf2,* and *Smg5* are all inviable (Metzstein and Krasnow, 2006; J.O. Nelson and M. M. Metzstein, *in prep*), and global overexpression of *Gadd45* leads to inviability (Peretz et al., 2007). Additionally, mosaic analysis has revealed that cells deficient for *Upf1* or *Upf2* in the egg chamber (Avery et al., 2011) or the eye (Metzstein and Krasnow, 2006) arrest growth in an

apoptosis-dependent manner (Avery et al., 2011), reminiscent of *Gadd45* overexpression in the egg chamber, which induces apoptosis (Peretz et al., 2007). Another striking phenotypic similarity manifests in eggshell patterning, where germline clones mutant for *Upf1* or *Upf2* result in fused eggshell dorsal appendages (Avery et al., 2011), a defect that can also be induced through overexpression of *Gadd45* in the germline (Peretz et al., 2007). The phenotypic similarities between NMD disruption and *Gadd45* overexpression are suggestive of an epistatic relationship and support our model that *Gadd45* could mediate a significant portion of the loss-of-function NMD phenotype.

Our conclusion that NMD mutants are physiologically affected by *Gadd45* upregulation is supported by our finding that disruption of *Mekk1* can also suppress the loss-of-function NMD phenotype. MTK1, as *Mekk1* is known in mammals, encodes a MAPKKK and is an obligate downstream signaling effector of *Gadd45.* MTK1 has been shown to be activated by GADD45 through direct protein/protein interactions, resulting in phosphorylation (Mita et al., 2002; Miyake et al., 2007; Takekawa and Saito, 1998). This phosphorylation activates MTK1, allowing it to act as an upstream regulator of MAPK signaling (Chen et al., 2002; Inoue et al., 2001), which in many contexts results in induction of apoptosis (Kang et al., 2012), similar to NMD loss of function mutations and Gadd45 overexpression.

Several other lines of evidence suggest a functional relationship between NMD, *Gadd45, Mekk1*, and MAPK signaling. While the direct activation of MEKK1 by GADD45 has not been tested in *Drosophila*, evidence suggests that activation of MAPK signaling could mediate the *Gadd45* overexpression phenotype. Disruption of *hemipterous* (*hep*), a signaling kinase protein upstream of jun N-terminal kinase (JNK)

and downstream of MEKK1, can rescue the fused dorsal appendage phenotype induced by Gadd45 overexpression (Peretz et al., 2007). Furthermore, in *Drosophila* S2 cell culture, *hep* is required downstream of *Mekk1* in the MAPK stress response pathway that is activated by cadmium and arsenite (Ryabinina et al., 2006).

*Gadd45/*MAPK signaling has also been studied in *Drosophila* in relation to stress responses that result from septic injury (De Gregorio, 2002; De Gregorio et al., 2001) and sterile wounding (Stramer et al., 2008) during which *Gadd45* is acutely upregulated. Other genes that are indicative of active JNK signaling, such as genes of the humoral immune system, are also upregulated during wounding. JNK pathway activation is functionally required for an effective response to wounding as JNK pathway mutants are cell-autonomously defective in wound healing (Galko and Krasnow, 2004; Rämet et al., 2002). The upstream kinases that initiate JNK signaling in wounding responses are not known; however, it is known that MEKK1 is activated during the wound response (Brun et al., 2006). For instance, the *Turandot* genes, which are involved in the humoral immune response (Ekengren and Hultmark, 2001), are strongly activated by wounding itself, and this response is blunted in animals mutant for *Mekk1* (Brun et al., 2006). We observed that many of these *Mekk1* dependent genes are also upregulated in $Upf2^{25G}/Y$ but are not reactivation targets (and thus upregulated indirectly), suggesting that *Mekk1* signaling is activated in NMD mutants.

Our data support a model where NMD is required to negatively regulate *Gadd45* throughout development (Figure 3.9). In NMD mutants, *Gadd45* transcript is stabilized, presumably resulting in increased level of GADD45 polypeptide and subsequent activation of MEKK1. Active MEKK1 in turn activates MAPK signaling pathways, such

as JNK and p38, resulting in transcriptional activation of downstream stress genes. We predict that when NMD mutant tissue succumbs to apoptosis, as occurs in the egg chamber (Avery et al., 2011), that this is initiated through stabilization of *Gadd45*. As precocious activation of *Gadd45* in the egg chamber during oogenesis causes widespread apoptosis and oocyte inviability (Peretz et al., 2007), it suggests that negative regulation of *Gadd45* by the NMD pathway begins at a very early time point in development. In support of this hypothesis, the creation of NMD mutant germline clones also results in apoptosis and oocyte inviability (Avery et al., 2011). To test the hypothesis that *Gadd45* stabilization results in apoptosis in NMD deficient germline tissue, germline clones could be generated that are double mutant for NMD genes and *Gadd45*, with the prediction that double mutant clonal tissue is resistant to apoptosis induced by loss of NMD pathway activity.

Our observations represent one example out of a very small number of instances where a loss of NMD function phenotype can be attributed to overexpression of a direct target. In *S. pombe*, double mutants between NMD genes and *rex2* (a direct target) can rescue the sensitivity to oxidative stress observed in NMD single mutants (Matia-González et al., 2013). In *Arabidopsis*, NMD activates an immune-like response (Rayson et al., 2012), and NMD defects can be ameliorated by introducing mutations that block full immune function (Riehs-Kearnan et al., 2012), suggesting that *Arabidopsis* NMD mutants succumb to autoimmune-like processes. As MAPK signaling is widely implicated as a positive activator of the *Drosophila* immune response, a very similar scenario may be at play in *Drosophila* NMD mutants.

The observation that $Gadd45^{54C}$ does not rescue NMD null mutants into

adulthood could suggest that there are additional targets of NMD and further suggests that NMD is required at multiple time points in development. For instance, *Gadd45* could be required after the L2/L3 molt, but this requirement is occluded due to *Gadd45*-independent lethality that arises before adulthood. However, our *Gadd45* allele is a mild disruption and could be indicative that this allele is hypomorphic in nature. Thus, full disruption of *Gadd45* function could result in a more complete rescue of animals harboring NMD null alleles. Therefore, it remains a possibility that *Gadd45* could solely be responsible for the NMD loss-of-function phenotype.

Our conclusions that *Gadd45* is directly and critically targeted by NMD raises questions about the role of the NMD-mediated type of regulation in development and physiology, specifically how NMD-mediated regulation may confer an evolutionary advantage. One possibility is related to the role of NMD in the silencing of viral transcripts. As NMD can efficiently target viral transcripts, many viruses have evolved countermeasures to evade such targeting. For example, the Rous sarcoma virus encodes *cis*-acting sequence features, known as RSV stability elements, that render its transcripts insensitive to NMD (Withers and Beemon, 2010). Additionally, the human T-lymphotropic virus type 1 (HTLV-1) expresses the protein Tax, which inhibits NMD in *trans* (Mocquet et al., 2012). These two observations imply that viruses are forced to overcome NMD surveillance to replicate in host cells. Moreover, cytomegalovirus vMIA protein has been shown to inhibit *Gadd45* (Smith and Mocarski, 2005). Taken together, these observations imply that NMD/*Gadd45* can act as a functional viral response pathway. One model concerning the role of negative regulation of *Gadd45* by NMD is that this system may have evolved to act as an early warning system to viral infection. If

a virus such as HTLV-1 enters a cell and inhibits NMD in *trans*, native target genes such as *Gadd45* become stabilized, thus activating MAPK/stress pathways shown to suppress infection.

Summary

In this chapter, a deficiency rescue screen is described that identified a potentially important NMD target gene, *Gadd45*. To directly test the role of this gene in the NMD pathway, TALEN technology was used to create the first partial loss-of-function alleles of *Gadd45*. These alleles partially rescue the L2 lethal phase of *Upf1* and *Upf2* null mutants. To our knowledge, this is the first time a functional, direct target of the NMD pathway has been identified. We also tested another gene identified in the screen, *Mekk1*, which is an obligate downstream effector of *Gadd45* activity, and showed that disruption of this gene also rescues the NMD loss-of-function phenotype. These results strongly suggest that *Gadd45*/*Mekk1* signaling is overactive in NMD mutants and contributes to inviability. The incomplete nature of the rescue phenotypes suggests that $Gadd45^{54C}$ is either a hypomorph and/or that other important direct NMD target genes exist and that NMD has many different roles that are necessary for viability and relies on different direct target genes to execute its function. It is likely that these other critical target genes exist in our list of 24 direct upregulated genes (Chapin et al., 2014).

Materials and methods

Stocks and reagents

Upf1^{26A}, *Upf1*^{13D}, *Upf2*^{25G}, and *Upf2*^{14J} are described in Metzstein and Krasnow (2006). For all experiments, these alleles are balanced over *FM7i-Act:GFP* for identification of hemizygous males in larval stages. Lethal phases were conducted using *Upf2*^{14J}*/FM7i; Gadd45*^{41B} and *Upf2*^{14J}*/FM7i; Gadd45*^{54C} females crossed to *FM7i/Y; Gadd45*^{41B} and *FM7i/Y; Gadd45*^{54C} males. For the *GMR:hid* analysis (Stowers and Schwarz, 1999), stocks of the following genotype were generated: *pGMR-hid, y*^{1}*, w** *FRT*^{19A}*/FM7c; ey-FLP*. These animals were crossed through stocks carrying the indicated deficiency before mated as males to stocks harboring NMD alleles. Final genotypes were *FRT*^{19A} *w* NMD*$^{-}$ */pGMR-hid, y*^{1}*, w** *FRT*^{19A} ; *ey-FLP/+; Df/+*. All genotypes were tracked using visual markers. For imaging, the eyes of anesthetized flies were imaged in several focal planes using a compound microscope and individual images compressed to a single, in-focus image using the stack focuser plugin for ImageJ (http://imagej.nih.gov/ij/). Next, ImageJ was used to calculate the area of each eye by manually defining the boarder of total ommatidia.

Cloning of the *eGFR:Gadd45* 3' UTR

To clone the 3' UTR of *Gadd45*, a 1050 bp fragment was PCR amplified from genomic DNA using Phusion polymerase and the following primers: G45_3U_X1_F and G45_3U_S1_R (Table 3.3). The template DNA used was from a stock that had been recently isogenized for the 2nd chromosome and harboring *FRT*^{G13}, the same stock used to induce *Gadd45* mutations. The fragment contained the 746 bp annotated 3' UTR of

Gadd45 plus 304 bases to ensure any cryptic polyadenylation signals were also included. This fragment was then TOPO-Blunt cloned into pCR4 and sequenced, revealing several SNPs when compared to the reference sequence published on FlyBase. This fragment was digested out of pCR4 using the *XbaI* and *StuI* sites included in the forward and reverse primers, respectively, and ligated into the "K1" vector (pUASTw/Attb-GFP) digested with Xba1 and Stu1. This procedure replaces the SV40 3' UTR with the newly cloned *Gadd45* 3' UTR. The final vector contains *UAS:eGFP-Gadd45* 3' UTR along with a backbone-localized *Attb* site for ΦC31-mediated transformation. The construct was injected into flies harboring the VK27 *Attp* docking site on the 3rd chromosome by Genetics Services, Inc. Three lines were established and shown to be functional by confirming GFP expression after crossing to *actP:GAL4* (not shown).

High-resolution melt analysis (HRMA)

Four primer pairs (all combinations of forward and reverse primers of: Gadd45_F3, Gadd45_F4, Gadd45_R7 and Gadd45_R8) were tested at three temperatures (68 °C, 70 °C, 72 °C) for optimal amplification of gDNA isolated from FRT^{G13}/+ animals. This led to the selection of Gad45_F3 and Gadd45_R7 (120 bp amplicon) at 70 °C for use in all downstream HRMA analysis. All G_0 and F_1 flies subjected to HRMA were homogenized in squishing buffer and 1ul of gDNA used in the PCR reaction. Following amplification, a HB96 light scanner (Idaho Technologies, UT) was used for the HRMA that would detect heteroduplex DNA and thus potential mosaic or heterozygous aberrations in the tested amplicon. Information about founding G_0 mutants is listed in Table 3.4.

Deficiency rescue screen

To mitigate the effects of background mutation, we selected a collection of stocks, referred to as the "DrosDel" collection, that carry custom, molecularly defined deficiencies that were generated in an isogenic w^{1118} background (Ryder et al., 2007). For our screen, we used 412 stocks from the DrosDel collection with deficiencies located on the 2nd, 3rd, and 4th chromosomes. *Upf2* is located on the X, and therefore our screen was not easily adaptable to studying any effect of deficiencies located on this chromosome. In total, this collection covered 67.4% of all bases (disrupting 63.7% of all loci) on the chromosomes tested (Table 3.5). To conduct the screen, males from each of the 412 stocks were individually crossed to $Upf2^{25G}/FM7i$ virgin females (Figure 3.1). Male F1 progeny from this cross were collected over a period of several days and scored for genotypes using visible markers. X chromosome genotypes could be distinguished by presence or absence of the *Bar* gene on *FM7i* balancing $Upf2^{25G}$. Chromosomes harboring the deficiencies could be distinguished by absence of the balancers in the Df stock (usually *CyO* for the 2nd and *TM6C, Sb* for the 3rd) or w^+ linked to the deficiency itself (Ryder et al., 2007).

To assess the impact of the various deficiencies on $Upf2^{25G}/Y$ viability, we measured the proportion of $Upf2^{25G}/Y$ to *FM7i/Y* between animals that harbored deficiencies and those that did not. To control for breeding conditions and the impact of the deficiency on overall (non-NMD dependent) viability, we compared the proportion of $Upf2^{25G}/Y$ males both carrying the deficiency and carrying the corresponding balancer within the same cross. The difference in the portion of $Upf2^{25G}/Y$ larvae that are balancer positive from the portion that are deficiencies positive gives a so-called Deficiency

Rescue Score (DRS) (Figure 3.1). A DRS of 0 indicates that the deficiency has a neural effect on $Upf2^{25G}/Y$ viability. Thus, a DRS of < 0 indicates the deficiency had a negative impact on viability, whereas a DRS > 0 indicates the deficiency has a rescuing effect. The majority of deficiencies tested (90.6%) had no significant effect on or were detrimental to $Upf2^{25G}/Y$ viability.

To generate mutant alleles of *Gadd45*

We used a TALEN-based approach that has been demonstrated to work in *Drosophila* (Liu et al., 2012) coupled to high resolution melt analysis (HRMA) for mutation detection. *Gadd45* is located on the right arm of the second chromosome, and we chose to induce mutations in a lab strain that carries FRT^{G13}, located at the base of 2R. This background was selected for a number of reasons: First, inducing mutations on the FRT-bearing chromosome would streamline subsequent experiments, such as mosaic analysis by eliminating the need to recombine the mutant alleles and the FRT. Second, FRT^{G13} is marked with w^+, making for a convenient linked marker with which to follow the mutant chromosome. Third, the FRT^{G13} chromosome in this stock had recently been isogenized procedure, assuring that recovered alleles should be in a similar genetic background.

Prior to TALEN design, Sanger sequencing was used to confirm that the locus encoding *Gadd45* in the FRT^{G13} stock matched the reference genome (data not shown). Next, TALEN monomers were designed to bind on either side of a spacer sequence between the 36th and 50th nucleotides of the predicted *Gadd45* CDS. *In vitro* transcribed mRNA coding for each TALEN was injected into G_0 embryos at either 200 or 400 pg/µl

and the embryos allowed to develop into adulthood (Table 3.4). Surviving G_0 adults (males and females) were mated with a w^- lab strain for several days before being isolated and homogenized to extract gDNA for HRMA analysis. In total, 21 fertile G_0s were screened from a total of 44 that survived the injection. w^+ F_1 progeny from the 14 G_0s that were identified as containing heteroduplex DNA (and therefore possible mutations in *Gadd45*) were selected for propagation. 5 through 30 F_1 progeny from each candidate G_0 were outcrossed to a CyO balancer stock and allowed to mate for several days before being removed and subjected to a similar HRMA procedure as the G_0s. This identified five candidate F1s lines carrying mutations in *Gadd45*. These five lines were propagated by sibling-pair matings between w^+; *Cy* F2s to create balanced lines. In all cases, Cy^+ progeny were observed in the F3 generation, indicating these putative mutations were not homozygous lethal.

Upon sequencing of homozygous F3s, 4 deletions were identified from 2 founding G_0 lines (Table 3.4). One of the original five lines, named *41B*, was revealed to be a false positive from the HRMA and had no changes from the original FRT^{G13} line. Three of the alleles, *41D, 54C,* and *54M,* are in-frame deletions of 18 bp, whereas *41E* eliminates 21 bp, and as such, none of these alleles are predicted to disrupt the *Gadd45* reading frame (Figure 3.5B). Very little is known about the structure/function of GADD45, it is therefore challenging to predict how these alleles will behave in comparison to wild-type. The amino acids deleted in these alleles are in an unconserved region of the protein compared to the three mammalian homologues (Figure 3.5C). We used $Gadd45^{54C}$ as a representative mutant allele and $Gadd45^{41B}$ as a control.

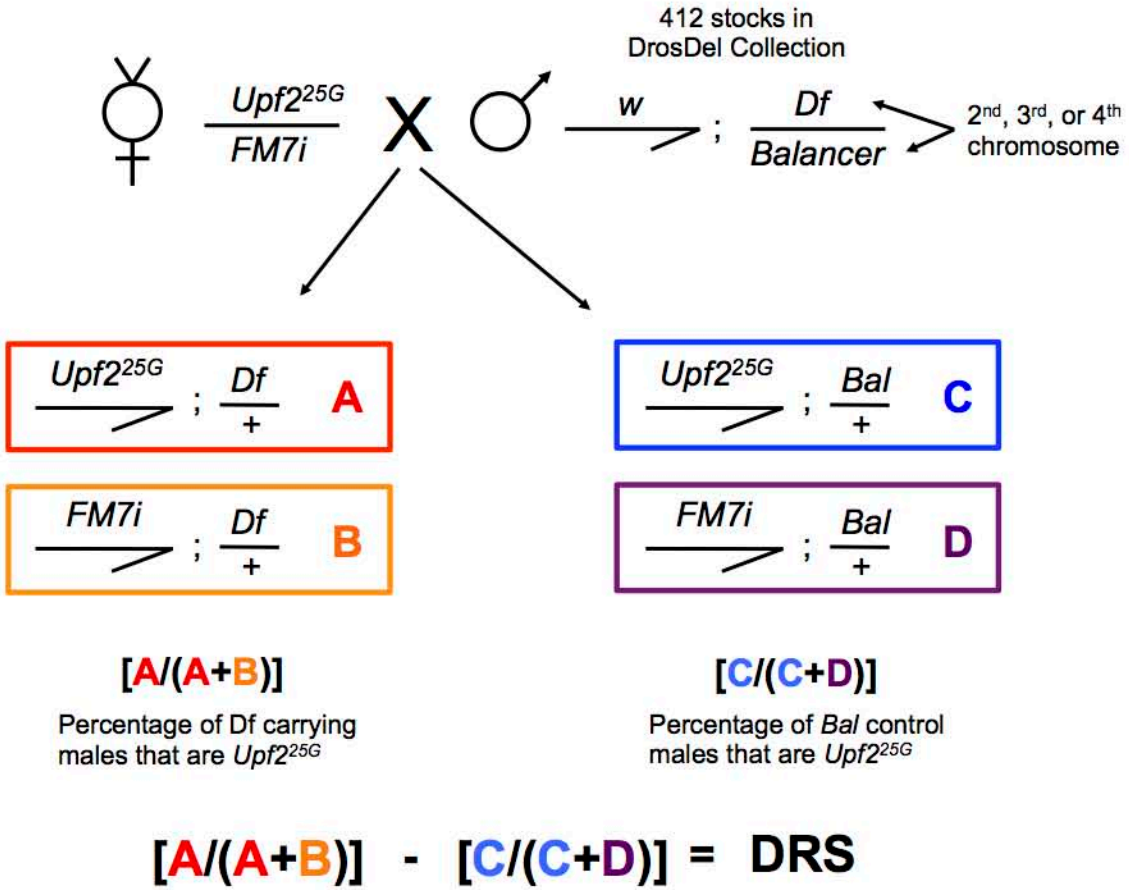


Figure 3.1 Strategy for genomic deficiency rescue screen. Males from 412, stocks part of the DrosDel collection, were crossed to *Upf2*^{25G}*/FM7i* females. Male F_1 progeny were scored for genotype using the *Bar* marker for the X chromosome and the appropriate marker corresponding to the autosomal balancer for each *Df* stock or the w^+ linked to the Df itself. The "deficiency rescue score" (DRS) is calculated as shown. The proportion of Balancer-carrying males that are *Upf2*^{25G}*/Y* is subtracted from proportion of *Df*-carrying males that are *Upf2*^{25G}*/Y*. This procedure controls for the presence of any background genetic changes in the deficiency-carrying stock that may effect viability, but are not linked to the deficiency itself.

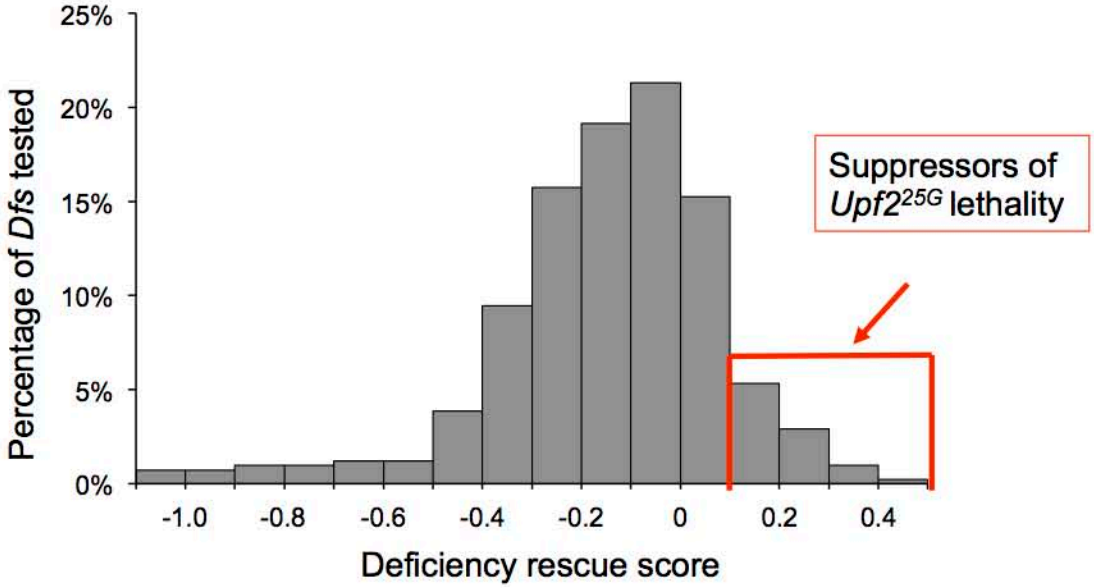


Figure 3.2 Histogram of deficiency rescue scores. Shown is the distribution of scores for all deficiencies in the collection. The histogram indicates the percentage of scores indicated on the x axis. The mean DRS is -1.4, and 9.4% of deficiencies had a DRS > 0.1, which is our threshold for rescue.

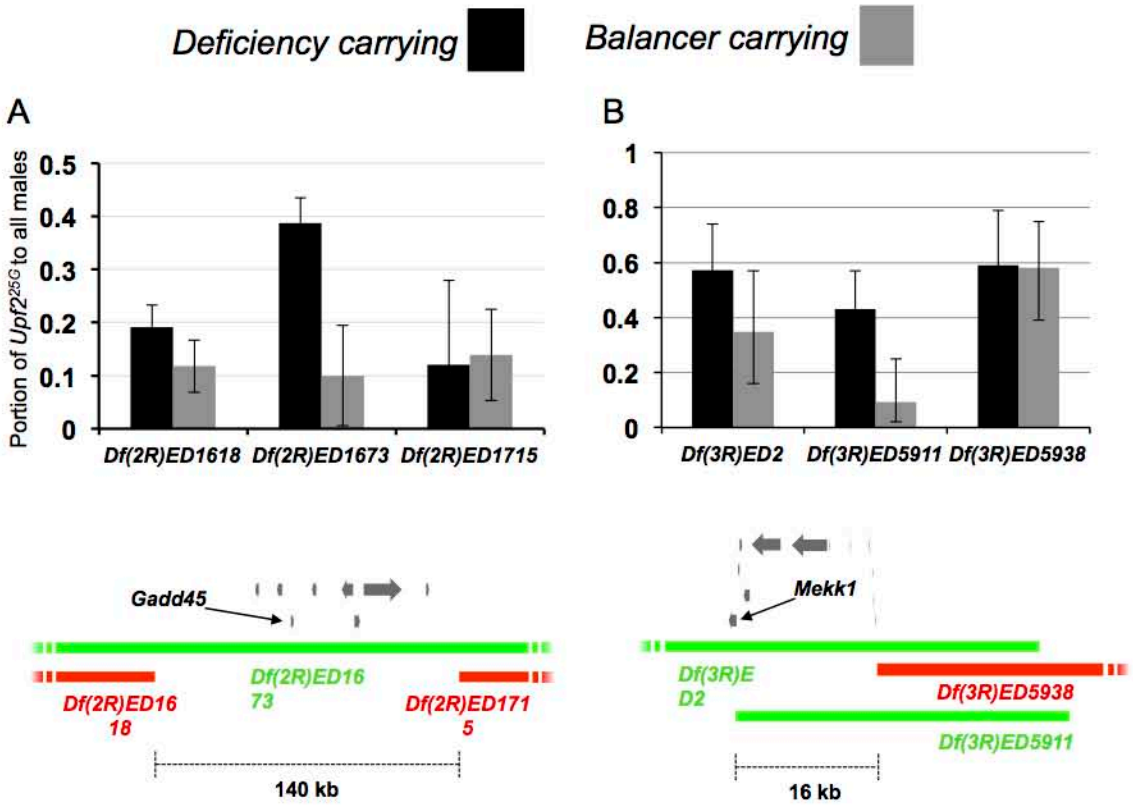
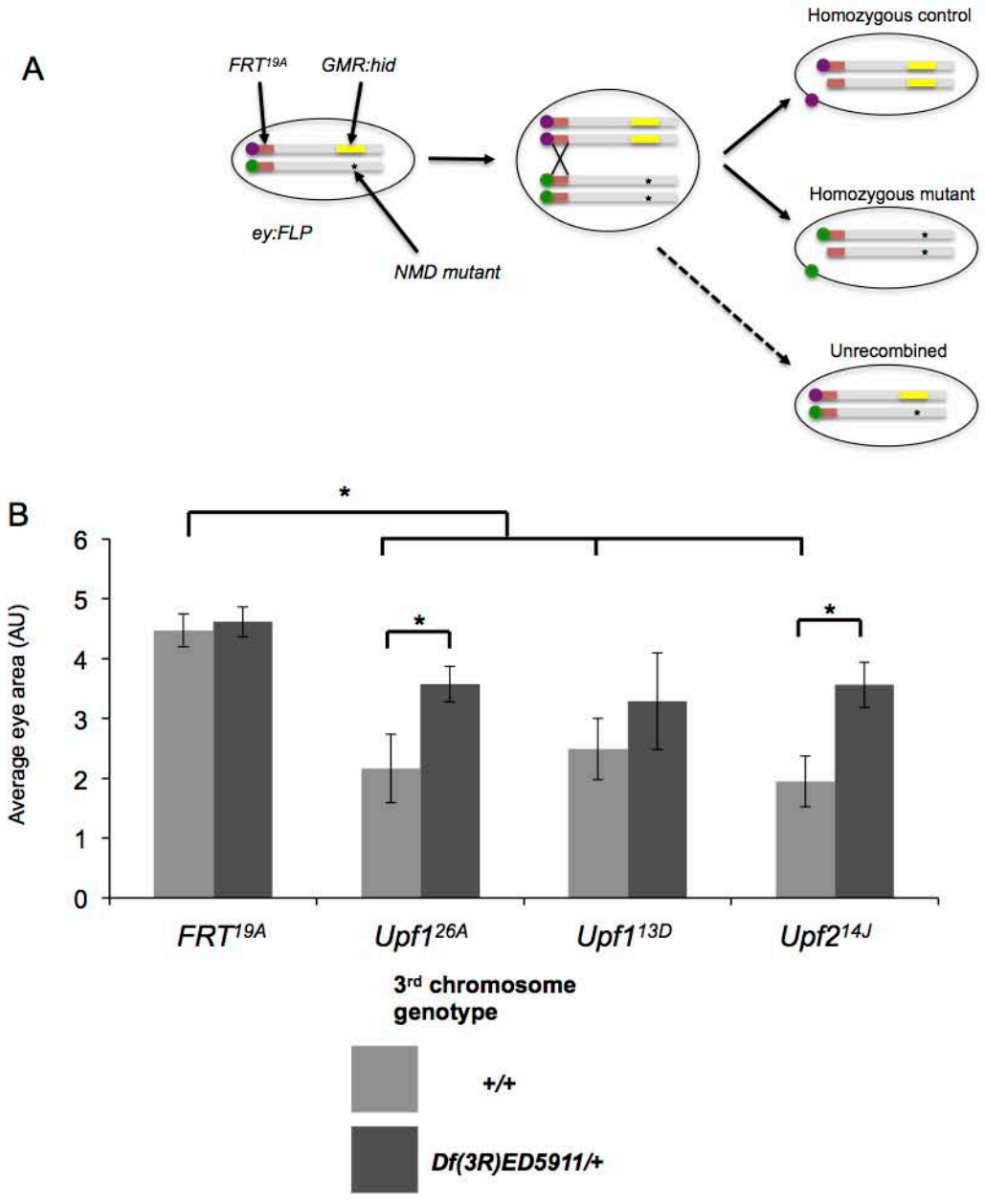


Figure 3.3 Deficiencies that uncover *Gadd45* and *Mekk1* rescue $Upf2^{25G}/Y$. Percentage of F1 $Upf2^{25G}/Y$ animals over total male animals from the deficiency rescue screen using the deficiencies indicated. Black bars indicate individuals harbored the deficiency, grey bars, the appropriate balancer. (A) *Df(2R)ED1618* and *Df(2R)ED1715* do not uncover *Gadd45*, whereas *Df(2R)ED1673* does. (B) *ED2* and *ED5911* both uncover *Mekk1*, whereas *ED5938* does not. Error bars indicated ± 95% confidence interval of the binomial. Maps show predicted transcriptional units in the neighborhood of the tested deficiencies.

Figure 3.4 Deficiencies that rescue *Upf2*^{25G}*/Y* subviability also rescue the small-eye phenotype revealed by the *GMR-hid* technique. (A) Schematic diagram indicating genetic recombination in eye precursor cells. Using this technique, mutant alleles of *Upf1* or *Upf2* are heterozygous over an X chromosome carrying a transgene that drives expression of the pro-apoptotic gene, *hid*, in the eye using the *GMR* promoter. FLP is driven by the *eyeless* promoter from a transgene on the second chromosome. When recombination is induced, NMD-mutant mosaic cells are generated such that the *GMR-hid* is segregated away, driving apoptosis in twin-spot and nonrecombined cells. (B) Total two-dimensional area of ommatidial tissue generated using the *GMR-hid* technique. Animals are anesthetized and the eyes imaged using multiple focal planes. Area of each eye is manually outlined in ImageJ and total area calculated. *FRT*^{19A} serves as a control for NMD alleles. Light grey bars indicate average eye size (arbitrary units (area)) with no deficiency present, dark grey bars, of flies carrying *Df(ED)5911,* a deficiency that rescues subviability in *Upf2*^{25G}*/Y* males. Error bars represent ± standard deviation. * indicates t-test value < 0.05.


A
FRT19A
GMR:hid
ey:FLP
NMD mutant
Homozygous control
Homozygous mutant
Unrecombined
B
Average eye area (AU)
0
1
2
3
4
5
6
*
*
*
FRT19A
Upf126A
Upf113D
Upf214J
3rd chromosome genotype
+/+
Df(3R)ED5911/+

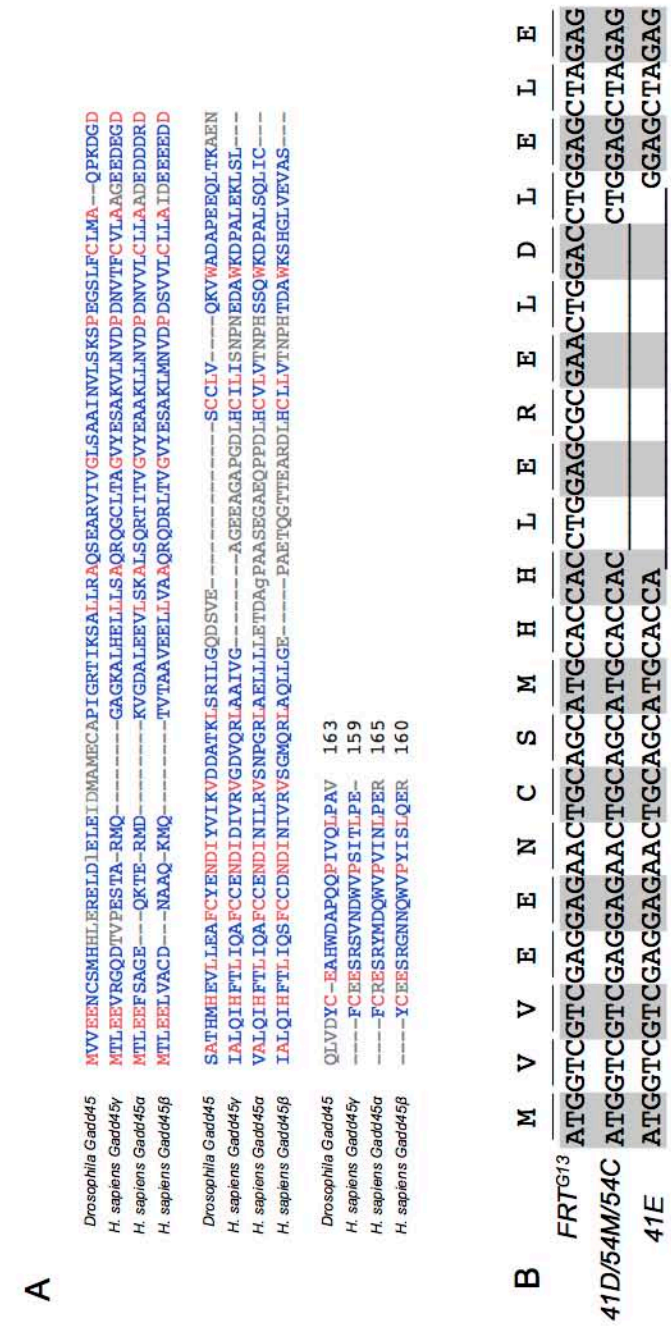


Figure 3.5 Alignment and allele generation of Gadd45. (A) Alignments of all three human homologues of *Gadd45* (top) and *Drosophila Gadd45* (bottom). (B) Protein alignments of recovered TALEN deletion mutants.

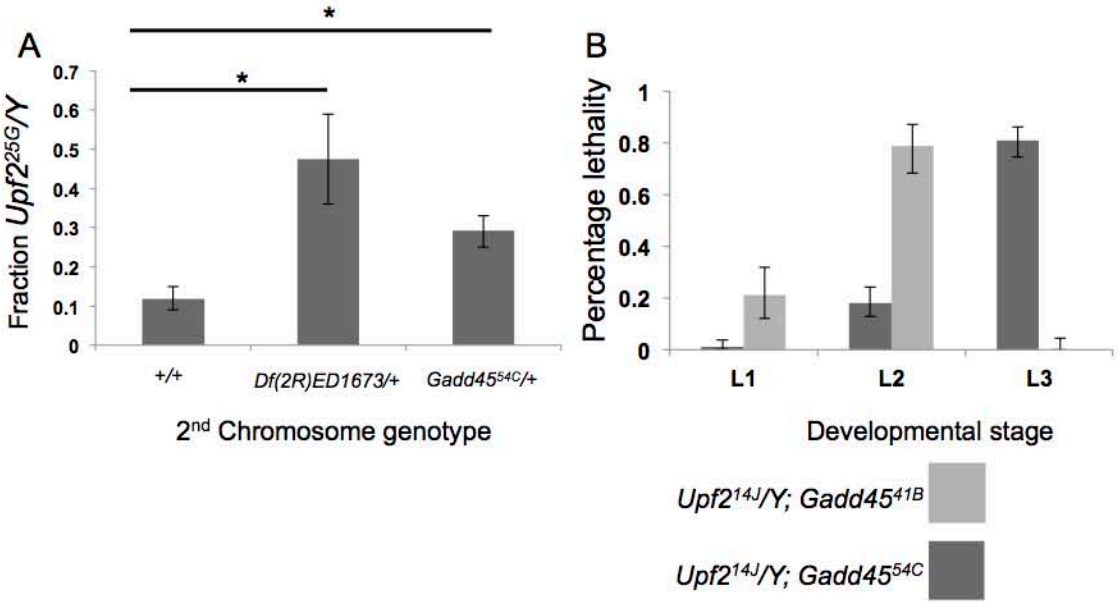


Figure 3.6 Rescue of NMD mutant defects by *Gadd45*^{54C}. (A) Percentage of F1 *Upf2*^{25G}*/Y* animals over *FM7i/Y* animals that are also the indicated genotype on the 2nd chromosome. *Gadd45*^{54C} heterozygotes are incompletely rescued compared to animals harboring *Df(2R)ED1673*. (B) The lethal phase of *Upf2*^{14J} larvae that are also homozygous for either *Gadd45*^{41B} or *Gadd45*^{54C}. Error bars represent $\pm$ 95% C.I. of the binomial. * equals chi-squared, p. value < 0.05.

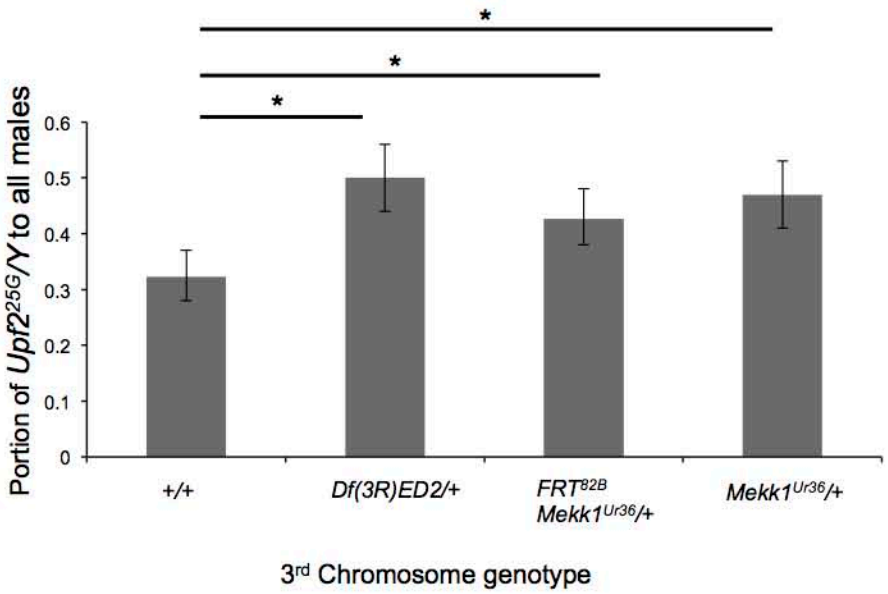


Figure 3.7 *Mekk1*^{Ur36} rescues *Upf2*^{25G}*/Y* animals. Percentage of F1 *Upf2*^{25G}*/Y* animals over total male progeny that are also the indicated genotype on the 3rd chromosome. Two versions of the *Mekk1*^{Ur36} chromosome were used, one harboring *FRT*^{82B} and one without any FRT. Error bars represent ± 95% C.I of the binomial. * equals chi-squared, p. value < 0.05.

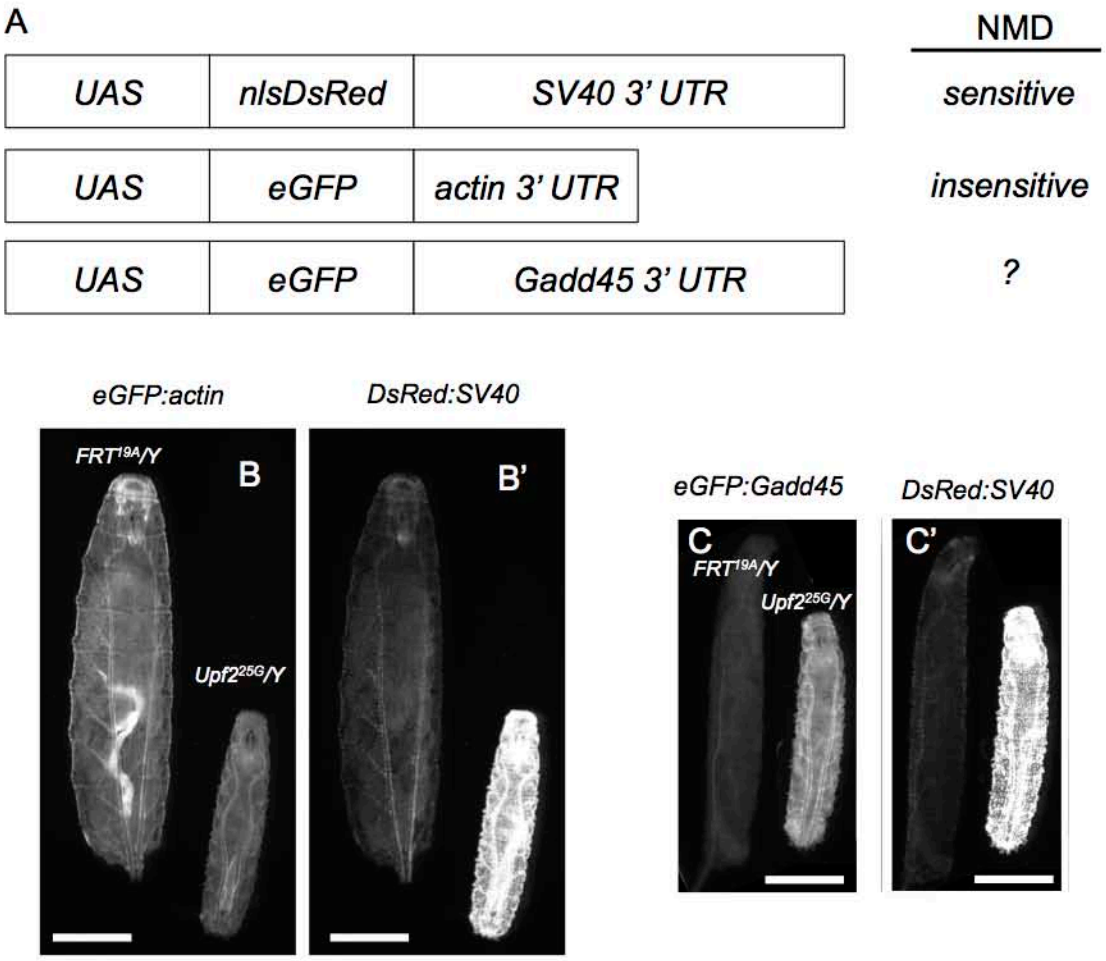


Figure 3.8 The 3' UTR of *Gadd45* is sufficient to render transcripts sensitive to NMD. (A) Schematic of transgenes generated in B. The *SV40* 3' UTR serves as an NMD-sensitive positive control and the *actin* 3' UTR as an NMD-insensitive negative control. (B, B', C, C') L3 larvae express an *actin:GAL4*-driven *UAS:nlsDsRed* transgene and either a *UAS:eGFP:actin 3' UTR* (B and B') or *UAS:eGFP:Gadd45 3' UTR* transgene (C and C'). In each panel, the larvae on the left is FRT^{19A}/Y and on the right is $Upf2^{25G}/Y$. Panels B and C show the green channel, revealing eGFP expression, and B' and C' show the red channel, revealing DsRed expression.

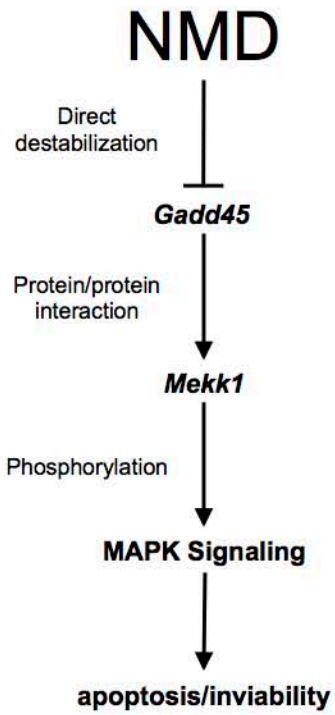


Figure 3.9 Model of NMD function.

Table 3.1 Deficiency coverage for the 24 upregulated direct targets identified in Chapin et al. (2014).

Gene Name	Number of deficiencies tested	DRS Scores
CG6908	0	NA
Xrp1	3	0.22, 0.33, 0.01
Ude	0	ND
CG13868	0	ND
CG34165	0	ND
Spat	0	ND
CG18748	2	0.08, 0.14
CG15784	0	ND
sra	0	ND
CG12766	1	-0.22
CG3008	1	-0.3
CG13659	0	ND
CG2906	0	ND
Fancl	1	-0.11
CG2017	2	-0.23, -0.07
CHKov2	0	ND
Smg6	1	-0.04
CG3635	0	ND
Smg5	0	ND
CG30424	0	ND
CG15661	1	-0.97
Ku80	0	ND
CG8620	0	ND
Gadd45	1	0.15

Table 3.2 Expression and reactivation data for *Mekk1*-regulated genes identified in Brun et al. (2006).

FBgn	Gene Name	Expression ratio in *Upf2*^{25G}/*Y*	Upregulation p value	Reactivation target
FBgn0028396	TotA	329.9	3.81E-66	na
FBgn0035176	CG13905	2.7	2.34E-05	na
FBgn0030929	CG15043	2.1	1.68E-03	0
FBgn0039593	CG9989	2.1	2.23E-02	0
FBgn0041182	TepII	2.0	9.51E-03	0
FBgn0035743	CG15829	1.5	1.54E-01	na
FBgn0031562	CG3604	1.0	8.20E-01	0
FBgn0031701	TotM	1.0	1.00E+00	0
FBgn0036587	CG4950	1.0	1.00E+00	0
FBgn0035412	CG14957	1.0	5.82E-01	0
FBgn0034296	CG10912	0.9	2.64E-01	0
FBgn0000477	DNaseII	0.8	2.70E-01	0
FBgn0032283	CG7296	0.6	1.51E-03	na

Table 3.3 Primers used in this study.

Primer Name	Sequence
G45_3U_X1_F	GAAGTCTAGA TAGGCGCGTGTGAGCGGGACAG
G45_3U_S1_R	gaag AGGCCT gttgataaaactgcatac
Gadd45_F3	AATATGGTCGTCGAGGAGAACTG
Gadd45_F4	CAAACGGATAACACCCACAATA
Gadd45_R7	AAGGGCCGACTTGATGGTG
Gadd45_R8	TGGCCATATCGATCTCTAGC

Table 3.4 Summary of the animals used in the HRMA analysis for allele recovery.

	RNA injected (pg/ul)		
	200	400	Total
Total injected	32	24	56
Alive	26	18	44
Fertile	11	14	25
Subjected to HRMA	10	11	21
WT Melt Curve	5	2	7
Mutant Melt Curve-low	5	4	9
Mutant Melt Curve-high	0	5	5

Table 3.5. Number of bases and number of genes uncovered by the deficiencies in our deficiency screen.

	Total Size		Deficiency coverage		Percentage	
Chromosome # and arm	Bases	Genes	Bases	Genes	Bases	Genes
X:	22424827	2330	0	0	0.0%	0.0%
2L:	23013543	2756	16146418	1885	70.2%	68.4%
2R:	21148708	3025	11030089	1417	52.2%	46.8%
3L:	24545557	2809	16442809	1776	67.0%	63.2%
3R:	27907053	3552	21764642	2661	78.0%	74.9%
4:	1353850	91	655159	47	48.4%	51.6%
Total:	120393538	14563	66039117	7786	54.9%	53.5%
Chromosomes tested:	97968711	12233	66039117	7786	67.4%	63.6%

References

Avery, P., Vicente-Crespo, M., Francis, D., Nashchekina, O., Alonso, C.R., and Palacios, I.M. (2011). Drosophila Upf1 and Upf2 loss of function inhibits cell growth and causes animal death in a Upf3-independent manner. Rna *17*, 624–638.

Azzalin, C.M., and Lingner, J. (2006). The double life of UPF1 in RNA and DNA stability pathways. Cell Cycle *5*, 1496–1498.

Behm-Ansmant, I., Gatfield, D., Rehwinkel, J., Hilgers, V., and Izaurralde, E. (2007). A conserved role for cytoplasmic poly(A)-binding protein 1 (PABPC1) in nonsense-mediated mRNA decay. Embo J. *26*, 1591–1601.

Brun, S., Vidal, S., Spellman, P., Takahashi, K., Tricoire, H., and Lemaitre, B. (2006). The MAPKKK Mekk1 regulates the expression of Turandot stress genes in response to septic injury in Drosophila. Genes Cells *11*, 397–407.

Cali, B.M., Kuchma, S.L., Latham, J., and Anderson, P. (1999). smg-7 is required for mRNA surveillance in Caenorhabditis elegans. Genetics *151*, 605–616.

Chapin, A., Hu, H., Rynearson, S.G., Hollien, J., Yandell, M., and Metzstein, M.M. (2014). In vivo determination of direct targets of the nonsense mediated decay pathway in Drosophila. G3 *4*, 485–496.

Chen, W., White, M.A., and Cobb, M.H. (2002). Stimulus-specific requirements for MAP3 kinases in activating the JNK pathway. J. Biol. Chem. *277*, 49105–49110.

De Gregorio, E. (2002). The Toll and Imd pathways are the major regulators of the immune response in Drosophila. Embo J. *21*, 2568–2579.

De Gregorio, E., Spellman, P.T., Rubin, G.M., and Lemaitre, B. (2001). Genome-wide analysis of the Drosophila immune response by using oligonucleotide microarrays. Proc. Natl. Acad. Sci. USA *98*, 12590–12595.

Ekengren, S., and Hultmark, D. (2001). A family of Turandot-related genes in the humoral stress response of Drosophila. Biochem. Biophys. Res. Commun. *284*, 998–1003.

Fornace, A.J., Alamo, I., and Hollander, M.C. (1988). DNA damage-inducible transcripts in mammalian cells. Proc. Natl. Acad. Sci. USA *85*, 8800–8804.

Frizzell, K.A., Rynearson, S.G., and Metzstein, M.M. (2012). Drosophila mutants show NMD pathway activity is reduced, but not eliminated, in the absence of Smg6. Rna *18*, 1475–1486.

Galko, M.J., and Krasnow, M.A. (2004). Cellular and genetic analysis of wound healing in Drosophila larvae. Plos Biol. *2*, E239.

Hodgkin, J., Papp, A., Pulak, R., Ambros, V., and Anderson, P. (1989). A new kind of

informational suppression in the nematode Caenorhabditis elegans. Genetics *123*, 301–313.

Hwang, J., and Maquat, L.E. (2011). Nonsense-mediated mRNA decay (NMD) in animal embryogenesis: to die or not to die, that is the question. Curr. Opin. Genet. Dev. *21*, 422–430.

Inoue, H., Tateno, M., Fujimura-Kamada, K., Takaesu, G., Adachi-Yamada, T., Ninomiya-Tsuji, J., Irie, K., Nishida, Y., and Matsumoto, K. (2001). A Drosophila MAPKKK, D-MEKK1, mediates stress responses through activation of p38 MAPK. Embo J. *20*, 5421–5430.

Jin, S., Tong, T., Fan, W., Fan, F., Antinore, M.J., Zhu, X., Mazzacurati, L., Li, X., Petrik, K.L., Rajasekaran, B., et al. (2002). GADD45-induced cell cycle G2-M arrest associates with altered subcellular distribution of cyclin B1 and is independent of p38 kinase activity. Oncogene *21*, 8696–8704.

Johansson, M.J.O., and Jacobson, A. (2010). Nonsense-mediated mRNA decay maintains translational fidelity by limiting magnesium uptake. Genes Dev. *24*, 1491–1495.

Kang, M.-J., Chung, J., and Ryoo, H.D. (2012). CDK5 and MEKK1 mediate pro-apoptotic signalling following endoplasmic reticulum stress in an autosomal dominant retinitis pigmentosa model. Nat. Cell Biol. *14*, 409–415.

Leeds, P., Wood, J.M., Lee, B.S., and Culbertson, M.R. (1992). Gene products that promote mRNA turnover in Saccharomyces cerevisiae. Mol. Cell Biol. *12*, 2165–2177.

Liu, J., Li, C., Yu, Z., Huang, P., Wu, H., Wei, C., Zhu, N., Shen, Y., Chen, Y., Zhang, B., et al. (2012). Efficient and specific modifications of the Drosophila genome by means of an easy TALEN strategy. J. Genet. Genomics *39*, 209–215.

Mak, S.K., and Kültz, D. (2004). Gadd45 proteins induce G2/M arrest and modulate apoptosis in kidney cells exposed to hyperosmotic stress. J. Biol. Chem. *279*, 39075–39084.

Matia-González, A.M., Hasan, A., Moe, G.H., Mata, J., and Rodríguez-Gabriel, M.A. (2013). Functional characterization of Upf1 targets in Schizosaccharomyces pombe. RNA Biol *10*, 1–9.

McIlwain, D.R., Pan, Q., Reilly, P.T., Elia, A.J., McCracken, S., Wakeham, A.C., Itie-Youten, A., Blencowe, B.J., and Mak, T.W. (2010). Smg1 is required for embryogenesis and regulates diverse genes via alternative splicing coupled to nonsense-mediated mRNA decay. Proc. Natl. Acad. Sci. USA *107*, 12186–12191.

Medghalchi, S.M., Frischmeyer, P.A., Mendell, J.T., Kelly, A.G., Lawler, A.M., and Dietz, H.C. (2001). Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. Hum. Mol. Genet. *10*, 99–105.

Metzstein, M.M., and Krasnow, M.A. (2006). Functions of the nonsense-mediated mRNA decay pathway in Drosophila development. PLoS Genet. *2*, e180.

Mita, H., Tsutsui, J., Takekawa, M., Witten, E.A., and Saito, H. (2002). Regulation of MTK1/MEKK4 kinase activity by its N-terminal autoinhibitory domain and GADD45 binding. Mol. Cell Biol. *22*, 4544–4555.

Miyake, Z., Takekawa, M., Ge, Q., and Saito, H. (2007). Activation of MTK1/MEKK4 by GADD45 through induced N-C dissociation and dimerization-mediated trans autophosphorylation of the MTK1 kinase domain. Mol. Cell Biol. *27*, 2765–2776.

Mocquet, V., Neusiedler, J., Rende, F., Cluet, D., Robin, J.-P., Terme, J.-M., Duc Dodon, M., Wittmann, J., Morris, C., Le Hir, H., et al. (2012). The human T-lymphotropic virus type 1 tax protein inhibits nonsense-mediated mRNA decay by interacting with INT6/EIF3E and UPF1. J. Virol. *86*, 7530–7543.

Neu-Yilik, G., Gehring, N., Hentze, M., and Kulozik, A. (2004). Nonsense-mediated mRNA decay: from vacuum cleaner to Swiss army knife. Genome Biol. *5*, 218.

Papathanasiou, M.A., Kerr, N.C., Robbins, J.H., McBride, O.W., Alamo, I., Barrett, S.F., Hickson, I.D., and Fornace, A.J. (1991). Induction by ionizing radiation of the gadd45 gene in cultured human cells: lack of mediation by protein kinase C. Mol. Cell Biol. *11*, 1009–1016.

Peretz, G., Bakhrat, A., and Abdu, U. (2007). Expression of the Drosophila melanogaster GADD45 homolog (CG11086) affects egg asymmetric development that is mediated by the c-Jun N-terminal kinase pathway. Genetics *177*, 1691–1702.

Rayson, S., Arciga-Reyes, L., Wootton, L., De Torres Zabala, M., Truman, W., Graham, N., Grant, M., and Davies, B. (2012). A role for nonsense-mediated mRNA decay in plants: pathogen responses are induced in Arabidopsis thaliana NMD mutants. PLoS ONE *7*, e31917.

Rämet, M., Lanot, R., Zachary, D., and Manfruelli, P. (2002). JNK signaling pathway is required for efficient wound healing in Drosophila. Dev. Biol. *241*, 145–156.

Rehwinkel, J., Letunic, I., Raes, J., Bork, P., and Izaurralde, E. (2005). Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. Rna *11*, 1530–1544.

Riehs-Kearnan, N., Gloggnitzer, J., Dekrout, B., Jonak, C., and Riha, K. (2012). Aberrant growth and lethality of Arabidopsis deficient in nonsense-mediated RNA decay factors is caused by autoimmune-like response. Nucleic Acids Res. *40*, 5615–5624.

Ryabinina, O.P., Suddian, E., and Iordanov, M.S. (2006). D-MEKK1, the Drosophila orthologue of mammalian MEKK4/MTK1, and Hemipterous/D-MKK7 mediate the activation of D-JNK by cadmium and arsenite in Schneider cells. BMC Cell Biology *7*, 7.

Ryder, E., Ashburner, M., Bautista-Llacer, R., Drummond, J., Webster, J., Johnson, G., Morley, T., Chan, Y., Blows, F., Coulson, D., et al. (2007). The DrosDel deletion collection: a Drosophila genomewide chromosomal deficiency resource. Genetics *177*, 615–629.

Smith, G.B., and Mocarski, E.S. (2005). Contribution of GADD45 family members to cell death suppression by cellular Bcl-xL and cytomegalovirus vMIA. J. Virol. *79*, 14923–14932.

Stowers, R.S., and Schwarz, T.L. (1999). A genetic method for generating Drosophila eyes composed exclusively of mitotic clones of a single genotype. Genetics *152*, 1631–1639.

Stramer, B., Winfield, M., Shaw, T., Millard, T.H., Woolner, S., and Martin, P. (2008). Gene induction following wounding of wild-type versus macrophage-deficient Drosophila embryos. EMBO Rep. *9*, 465–471.

Takekawa, M., and Saito, H. (1998). A family of stress-inducible GADD45-like proteins mediate activation of the stress-responsive MTK1/MEKK4 MAPKKK. Cell *95*, 521–530.

Viegas, M.H., Gehring, N.H., Breit, S., Hentze, M.W., and Kulozik, A.E. (2007). The abundance of RNPS1, a protein component of the exon junction complex, can determine the variability in efficiency of the Nonsense Mediated Decay pathway. Nucleic Acids Res. *35*, 4542–4551.

Weischenfeldt, J., Damgaard, I., Bryder, D., Theilgaard-Mönch, K., Thoren, L.A., Nielsen, F.C., Jacobsen, S.E.W., Nerlov, C., and Porse, B.T. (2008). NMD is essential for hematopoietic stem and progenitor cells and for eliminating by-products of programmed DNA rearrangements. Genes Dev. *22*, 1381–1396.

Withers, J.B., and Beemon, K.L. (2010). Structural features in the Rous sarcoma virus RNA stability element are necessary for sensing the correct termination codon. Retrovirology *7*, 65.

Wittkopp, N., Huntzinger, E., Weiler, C., Saulière, J., Schmidt, S., Sonawane, M., and Izaurralde, E. (2009). Nonsense-mediated mRNA decay effectors are essential for zebrafish embryonic development and survival. Mol. Cell Biol. *29*, 3517–3528.

Yoine, M., Nishii, T., and Nakamura, K. (2006). Arabidopsis UPF1 RNA helicase for nonsense-mediated mRNA decay is involved in seed size control and is essential for growth. Plant Cell Physiol. *47*, 572–580.

CHAPTER 4

EVIDENCE OF SPECIFC SPATIAL AND TEMPORAL FUNCTIONS OF THE NMD PATHWAY

Introduction

The developmental requirement for NMD components indicates that this pathway has functions beyond the targeting and decay of nonsense alleles. The "Swiss Army Knife" model of NMD function states that NMD directly regulates a handful of critical target genes in a way that is essential for viability (Neu-Yilik et al., 2004). This model is in opposition to the "vacuum cleaner" model, which states the critical function of NMD is related to its role as a surveillance pathway degrading nonfunctional RNAs. One prediction of the Swiss Army Knife model is that NMD may have tissue-specific functions as expression patterns of critically regulated native target genes could be tissue-specific. Similarly, this model also states that NMD could have temporal-specific critical functions. For example, NMD may critically target a specific transcript in particular tissue and/or at a particular time in development in a way that is sufficient for organismal viability. Therefore, if tissues can be identified where NMD is either sufficient or necessary for viability, it would suggest not only that essential and specific target genes exist, but also indicate in what tissues such critical targets are expressed. Furthermore, by

experimentally controlling when NMD is active over the course of development, we can identify any temporal specificity of NMD function that may exist.

To test for critical spatiotemporal roles of NMD, we conducted a tissue-specific rescue screen for tissues where NMD function is sufficient for viability. This screen revealed that function of NMD neuronal tissue is sufficient to rescue the L2/L3 molting defect of *Upf2* null mutants. Additionally, by employing a genetic-based method to control when NMD components are active over the course of development, we find that function of zygotic NMD components during embryogenesis and midlarval stages may be sufficient to rescue NMD null mutants into adulthood. Taken together, these results support the Swiss Army Knife model, suggesting that NMD has specific, critical functions in neuronal tissue early in development and that NMD may be dispensable for viability in all adult tissues.

Results

Tissue-specific rescue screen

In this screen, a library of spatially restricted GAL4 drivers were used to restore NMD function in an NMD-defective genetic background. $Upf2^{14J}$ was used as representative NMD-defective background and therefore $UAS{:}Upf2^{WT}$ was used to rescue function. 80% of $Upf2^{14J}/Y$ larvae die as second instar larvae (Chapin et al., 2014), and we defined "rescue" as development to wandering third larval instar (WL3). The GAL4 driver library consisted of a collection of 30 stocks available from the Bloomington Stock Center that we characterized as having a restricted expression pattern along with three ubiquitous expressing drivers (Materials and Methods; Table 4.1). *FM7i-act:GFP/Y;*

GAL4/+ animals were crossed to *Upf2^{14J}/FM7i; UAS:Upf2^{WT}* (Figure 4.1A). By examining GFP expression provided by the *FM7i-act:GFP,* mutant (GFP negative) and nonmutant (GFP positive) WL3 progeny could be scored in the F2. 1/4 of total progeny in this cross are *Upf2^{14J}/Y; GAL4/UAS:Upf2* or *Upf2^{14J}/Y*; *+/UAS:Upf2*. As larvae of the latter genotype die in the second larval stage (L2), the presence of GFP-negative WL3 at a frequency of 1/7 total progeny indicates 100% rescue.

To assess the rate at which escaper *Upf2^{14J}/Y; +/UAS:Upf2* larvae develop into WL3, we first examined the calculated degree of rescue when no *GAL4* is present. This proved to be very minimal under the breeding conditions used; 3% of expected (Figure 1B). Next, we confirmed that the *UAS:Upf2^{WT}* transgene can rescue to WL3 when it is globally expressed using an *Act5C:GAL4* driver. We observed that this *GAL4* expression rescues to 138% of expected for full rescue (Figure 4.1B). That fact this rescue is greater than 100% is not unexpected as animals carrying the *FM7i* balancer are slightly developmentally delayed compared to nonbalancer, wild-type animals (data not shown), implying that our rescued *Upf2^{14J}* mutants are actually more fit that balancer-carrying larvae. Similar results were obtained when using other ubiquitous drivers (Figure 4.1B).

Most GAL4 drivers did not exhibit any capacity to rescue. However, a handful of drivers with restricted expression patterns rescue to varying degrees (Figure 4.1B). For instance, *P{GawB}48Y* was able to rescue to 113% of predicted. This driver expresses in the central nervous system (CNS) and imaginal discs, in addition to various cells in the trachea, midgut, and hindgut. *P{GAL4-Hsp70.PB}31-1* rescues to a lesser degree (64%) but also has a more restricted pattern, limited to the CNS, midgut, and hindgut. Lastly, *P{GawB}60IIA* rescues to 50% and is limited in expression to neuronal tissue and weak

expression in the epidermis and imaginal discs. We also observed partial rescue with *P{GawB}insc*^{Mz1407} (39%) and *P{GAL4-elav.L}2* (43%), both of which have expression in the CNS. One unifying feature of the expression patterns of these five GAL4 drives is expression in the CNS. In particular, *P{GawB}60IIA* is quite restricted to and strongly expressed in neuronal tissue (Figure 4.2). These data suggest that NMD activity in the CNS is sufficient for development into the WL3 stage.

NMD activity during late embryogenesis and larval stages may be sufficient for adult viability.

To test the hypothesis that the critical activities of NMD occur at limited time points in development, we borrowed logic from classical experiments that define critical periods of gene function through characterization of the "temperature-sensitive period" of temperature-sensitive alleles (Tarasoff and Suzuki, 1970). These classic experiments rely on activating and deactivating gene function by using temperature sensitive alleles. As no such alleles of NMD components are available in fly, we used *Act5C:GAL4* combined with a temperature-sensitive *GAL80* to control expression of a rescuing *UAS:Upf2* cDNA in *Upf2*^{14J} males (Figure 4.3A). As this GAL80 unfolds and loses activity at 29 °C, this temperature is permissive for NMD function by allowing for *GAL4*-mediated expression of the rescuing *UAS:Upf2*. At the restrictive temperature of 18° C, GAL80 folds appropriately, thus inhibiting expression of *UAS:Upf2*.

We first ensured that expression of *UAS:Upf2* could be modulated using temperature in a way that is physiologically relevant. To do this, we measured the proportion of *Upf2*^{14J}*/Y* males to total male progeny that developed when held

constitutively at the permissive temperature (29 °C) or the restrictive temperature (18 °C). This analysis revealed that *Upf2*^{14J}*/Y* adult males developed to 53% of expected at 29 °C (Figure 4.3B). In separate experiments where no GAL80 is present, *UAS:Upf2* driven by actin GAL4 rescues *Upf2*^{14J}*/Y* to 100% of expected (data not shown), indicating that the *GAL80*^{ts} used in the constitutively permissive experiment may have some residual activity at 29 °C. When *Upf2*^{14J}*/Y; Act5C:GAL4/+, tubP:GAL80*^{ts}*/+* larvae are held at the restrictive temperature, no adults enclosed, indicating that *GAL80* was able to inhibit expression of the *UAS:Upf2* cDNA to levels incompatible with viability (Figure 4.3B). Note that this experimental approach cannot be used to inhibit activity of maternally deposited, wild-type *Upf2* gene product. Thus, this experiment can only assay the zygotic requirements for *Upf2*.

To test for the end point of any NMD critical period, we tested for the latest time point larvae could be shifted from the permissive to restrictive temperate and still allow for development into adulthood at rates equal to that of constrictive NMD function at the permissive temperature. Multiple parental broods (Materials and Methods) were bred in vials at 29 °C, and F1 progeny allowed to develop at this temperature before being downshifted to 18 °C at various times points 0−6 days after egg lay. Adult F1s were collected, and the portion of *Upf2*^{14J}*/Y* males out of total progeny was calculated for each time point representing days before the downshift (Figure 4.3C, red line). From this analysis, we find that NMD activity can be inactivated as early as 2 days after egg lay and still allow for a small portion of *Upf2*^{14J}*/Y* to development into adulthood. When NMD is inactivated at 4 days after egg lay, a portion of *Upf2*^{14J}*/Y* males eclose which is similar to that of our control experiment where animals are kept at the permissive temperature

throughout development. These results suggest that zygotic *Upf2* is only required at early time points in development, as opposed to being strictly required for steady state physiology and that this critical window ends sometime in late embryogenesis/early larval stages.

To identify the beginning of this critical window of NMD function, larvae were bred at the restrictive temperature and upshifted to the permissive temperature at time points identical to those in the downshift experiment. We find that even very early activation of Upf2 does not lead to full viability of $Upf2^{14J}/Y$, suggesting that zygotic *Upf2* is required at these early time points (Figure 4.3C, blue line). Taken together, our upshift and downshift data suggest that the critical period of zygotic NMD function is from less than 6 hours after egg lay to less than 4 days after egg lay, roughly corresponding to embryogenesis and midlarval stages.

Functional assessment of RNAi-mediated silencing of NMD components

In the future, a tissue-specific RNAi screen could be used to assess the necessity of NMD components in various tissue types. A screen such as this would likely utilize the same library of GAL4 drivers in combination with *UAS* constructs that induce inhibition of NMD pathway components. As a proof of principle experiment, we tested several *UAS* transgenes designed to induce RNA-mediated silencing of NMD components, or expression of a dominant negative form of *Upf1* (UPF1-DN), for their ability to inhibit NMD. RNAi transgenes for all NMD components were publically available, except Upf1, for which we generated using two constructs targeting separate regions of *Upf1* (see

Materials and Methods).

To test the impact of these NMD-inhibiting transgenes on NMD function, all RNAi/DN stocks were crossed to a reporter stock that carried *btl:GAL4* and an NMD-sensitive *UAS:eGFP-SV40 3' UTR* reporter (Metzstein and Krasnow, 2006). Tracheal GFP expression in the third larval stage was compared between F1 progeny both with and without the RNAi or DN constructs. In all cases, targeting of any NMD factor with any of the RNAi/DN transgenes produced an upregulation in eGFP signal (Figure 4.4). As NMD has been shown to have no effect on transcription of GAL4 (Metzstein and Krasnow, 2006), our data indicate that all these NMD-inhibiting transgenes have a stabilizing effect on the *eGFP* RNA, indicating NMD function as at least partially compromised.

Discussion and future directions

Function of the NMD pathway is predicted to have critical roles in development and physiology. However, especially in higher eukaryotes, these roles are very poorly understood, and it remains unknown how the NMD pathway might fit into regulation of these processes. At the molecular level, NMD has been shown to function as a surveillance pathway and to be involved in the regulation of specific target genes, but it is not known which types of target are regulated most critically. The data presented in this chapter support the Swiss Army Knife model of NMD function, such that, at least in *Drosophila*, NMD critically regulates specific target genes. First, as NMD is thought to function in most tissue types (Avery et al., 2011; Metzstein and Krasnow, 2006; Rehwinkel et al., 2005), our results that an NMD null mutant can be partially rescued in a tissue-specific manner imply that the physiologically relevant targets of NMD are also

tissue-specific. Furthermore, our experiments to identify critical time periods of NMD function suggest that function in early embryogenesis or early larval stages is sufficient for viability into adulthood. As aberrant or "junk" transcripts are thought to be produced at all times in all tissues throughout development, our observation that NMD pathway function is not required in the adult suggests that the surveillance capabilities of NMD are not necessary for viability.

Taken together, our data support a model where NMD function in the central nervous system (CNS) during embryogenesis and early larval stages is sufficient for viability into adulthood. This model is consistent with evidence from *Drosophila* and mammalian studies that NMD functions in important postmitotic and development processes in neuronal tissue, such as in long-term potentiation at synapses (Giorgi et al., 2007), proper synaptic vesicle cycling and morphology of the neuromuscular junction (Long et al., 2010), axon guidance (Colak et al., 2013), and proper brain development and/or function in human (Tarpey et al., 2007). Also in agreement with our model are observations in mice that NMD activity is gradually downregulated in neuronal tissue over the course of development and that this downregulation is actually required for neurodevelopment (Bruno et al., 2011). If NMD is regulated in a similar way in *Drosophila*, it would agree with our data suggesting NMD activity is not required in adults. Moreover, this model predicts that overactivation of NMD in *Drosophila* would inhibit neurodevelopment, or perhaps produce other gain-of-function NMD phenotypes. If overactivation of NMD does have detrimental effects in *Drosophila*, it could explain observations from our lab that overexpression of *Upf1*, even when using tissue-restricted GAL4 drivers, often leads to lethality. However, it is not known if overexpression of

Upf1 in fact leads to hyper activation of the NMD pathway.

One possible mechanism by which NMD activity in the CNS controls the L2/L3 molt is through regulation of the mechanisms that result in production of 20-hydroxyecdysone (20E), a steroid hormone that controls molting and pupariation (Edgar, 2006; Metzstein and Krasnow, 2006). Under wild-type, nutrient-rich conditions, ecdysone (a precursor to 20E) is produced and released by the prothoracic gland in response to stimulation by prothoracicotropic hormone (PTTH). PTTH is synthesized and released in a highly regulated fashion by a pair of neurosecretory cells in the CNS that innervate the prothoracic gland (McBrayer et al., 2007). Given our rescue data, it could be possible that the NMD is required in some process that stimulates PTTH release and by extension, the initiation of molting/pupariation. This model could be tested by asking if *ptth:GAL4* (McBrayer et al., 2007), which expresses only in the PTTH-producing cells, can rescue the L2/L3 the molting defect in $Upf2^{14J}$.

In one model, NMD mutants upregulate putative negative regulators of PTTH release, and restoration of NMD in neuronal tissue alone renormalizes these negative regulators to wild-type levels, allowing for normal PTTH release and thus rescue of the L2/L3 molt. However, this model is based on necessity of NMD pathway activity in the CNS, which our tissue-specific rescue screen did not directly test. To examine the necessity of NMD in CNS, it will be important to test whether tissue-specific inhibition of NMD in the CNS is incompatible with developmental progression. Such an experiment could be executed using RNAi transgenic reagents combined with our rescuing GAL4 drivers, provided that it can be demonstrated that these RNAi/DN transgenes can induce physiologically relevant silencing of NMD components. Several

attempts were made to use these RNA/DN constructs to study neuron-specific aspects of the NMD loss of function phenotype. However, these results were inconclusive for many reasons. First, many transgenes designed to silence NMD components previously characterized as vital did not induce lethality when expressed globally using *Act5C:GAL4*. In the future, techniques aimed at enhancing the ability of RNAi transgenes to induce silencing, such as co-expression with siRNA pathway component DICER, or incubating animals at higher temperatures, could be used to augment the silencing ability of the available RNAi lines. In pilot studies, the presence or absence of visible markers was used to assess whether the transgenic combinations were lethal, and any RNAi or GAL4 transgenes that were mapped incorrectly would cause this approach to fail. Thus, these transgenic stocks should be rebalanced to ensure that all visible markers correspond to the appropriate genotype.

A second, different interpretation of our CNS rescue data is that NMD function in the CNS is sufficient, but not necessary for developmental progression. The interruption of the L2/L3 molt that occurs in $Upf2^{14J}$ could be caused by defects that arise in nonneuronal tissues, yet activation or overexpression of *Upf2* in the CNS bypasses these nonneuronal defects and initiates the L2/L3 molt regardless. Interestingly, this model fits with the current understanding of PTTH-mediated stimulation of ecdysone release. It has been observed that ablation of the PTTH producing cells does not lead to full inhibition of molting or pupariation, suggesting that functional ecdysone release can be stimulated through other channels (McBrayer et al., 2007). If it is found that NMD function in the CNS is not necessary for viability, one interpretation would be that although PTTH release is defective in animals with CNS-specific silencing of NMD components, NMD

activity in nonneuronal tissue is sufficient to induce ecdysone release. When NMD mutant animals are rescued with expression in CNS, while lack of NMD activity in nonneuronal tissues is not sufficient to drive molting, expression in the CNS is sufficient for PTTH release.

This proposed mechanism would also fit with our data showing that $Upf1^{26A}$ and $Upf2^{14J}$ mutants developmentally arrest for weeks before they die as second or third instar larvae (data not shown). These data suggest that loss of NMD pathway function is not acutely lethal to larvae, but rather inconsistent with developmental progression. If this arrest is related to insufficient ecdysone production, it would be possible under this proposed mechanism that the induction of pathways leading to ecdysone release would rescue the molting defects of *Upf1* and *Upf2* null mutants. A test of this model would be to treat NMD mutants with purified 20E and ask whether this is also sufficient to bypass the second instar arrest.

Our interpretation that NMD function in the CNS is sufficient to rescue the L2/L3 molting defect in NMD mutants does have caveats. In the future, it will be important to ensure that NMD function has been restored only to the tissues were we believe the GAL4 to be expressed. Such an experiment could be conducted using expression of NMD-sensitive reporters in NMD-rescued larvae. Expression patterns of these GAL4s are not well characterized beyond our own work observing expression of fluorescent reporters (Table 4.1) and could have functional expression in other tissues that we can not observe using such reporters. Other important controls for this sufficiency experiment would include using a separate *UAS:Upf2* cDNA to control for positional effects of individual transgenes. It would also be important to demonstrate rescue of other null

alleles of *Upf2* and perhaps Upf1 (which also arrest at the L2/L3 molt), although this latter experiment might be impossible as overexpression of *Upf1* in a number of specific tissue types results in organismal lethality.

While fitting with certain data, our conclusion that the critical period of NMD function ends in larval stages could be at odds with other observations concerning NMD function in the fly. For instance, as has been shown in cell culture (Rehwinkel et al., 2005) and whole *Drosophila* (Avery et al., 2011; Metzstein and Krasnow, 2006), NMD function is required for cell proliferation. As cell proliferation is required in pupal stages, it is interesting that we identified the end of the critical period as between 3 and 4 days, a time point corresponding to early L3 and over 24 hours before pupariation. Additionally, *Smg5* null mutants die during pupariation (J.O. Nelson and M. M. Metzstein, *in prep*), which also suggest that NMD has critical functions at this stage. This disparity could be explained in two ways. First, our temperature sensitive period experiments are designed to identify critical periods of gene function, not to identify when mutant defects manifest. Therefore, it is possible that some aspect of NMD function during early developmental time points is important for the cell proliferation that occurs during pupal stages. Another explanation for this disparity is related to the limitations of our experimental design. It is possible that, while the shift to 18 °C does inhibit GAL80, the perdurance of the RNA or protein corresponding to the *Upf2* cDNA lasts into pupal stages, thus rescuing any critical NMD-mediated functions at these later time points. If this perdurance is long enough, it could also explain why we observe any adult *Upf214J/Y* males at all in the downshift experiment to the restrictive temperature.

The issue of perdurance is a vitally important caveat to the interpretation of these

data. When this experiment is repeated in the future, measurements should be made of *Upf2* mRNA and protein levels throughout the downshift time course. Additionally, NMD function itself, as measured through upregulation of direct NMD target transcripts, should also be assayed at various time points to ensure that NMD is in fact inhibited in a timely manner following the shift to the restrictive temperature. If it cannot be demonstrated that NMD is sufficiently inhibited in pupae and adults following downshift in larvae, it will invalidate our conclusions about this critical period of NMD function.

Materials and methods

Stocks and reagents

Upf1^{26A} and *Upf2*^{14J} are as described in Metzstein and Krasnow (2006). For the "temperature sensitive period" experiments, the GAL80ts is *P{w[+mC]=tubP-GAL80[ts]}10* (from Bloomington Stock Number 1708) recombined with *Act5C: GAL4*. The *UAS:Upf2* is as described in Chapin et al., (2006). The *GAL4* reporter line harbors *UAS:GFP* and *UAS:nlsDsRed* reporters. A complete list of the lines in the tissue-restricted GAL4 library is available in Table 4.1.

Cloning of *Upf1* RNAi transgene

To clone RNAi constructs directed against *Upf1*, two target sequences were cloned, an "A" and a "B" of 437 and 422 base pairs, respectively. Each target sequence was amplified from gDNA two times, once using primer pairs that create a 5' *NotI* site and a 3' *XhoI* site ("left") and again using primer pairs that create a 5' *XbaI* site and a 3' KpnI site ("right"). Primers sequences are available in Table 4.2. All 4 fragments were

first TOPO-TA cloned into PCR4 (Invitrogen, CA) and then sequentially ligated into pUAST-I w/Attb starting with the "left" fragments for both the "A" and "B" versions. This vector is designed to position the "left" version of the target sequence upstream of the "right" version, such that the "right" and "left" are co-expressed as a sense and antisense version, creating double stranded RNA and thus inducing RNAi directing towards the target sequence. Following completion of the cloning, both plasmids were sequenced and shown to contain no changes from the published sequence. These constructs were injected into flies harboring the VK26 (3^{rd} chromosome) or Attp16 (2^{nd} chromosome) AttP sites for ΦC31-mediated integration. Transgenic stocks were isolated by virtue of expression of the w^{+} minigene in pUAST-I and balanced with TM3, Sb or CyO.

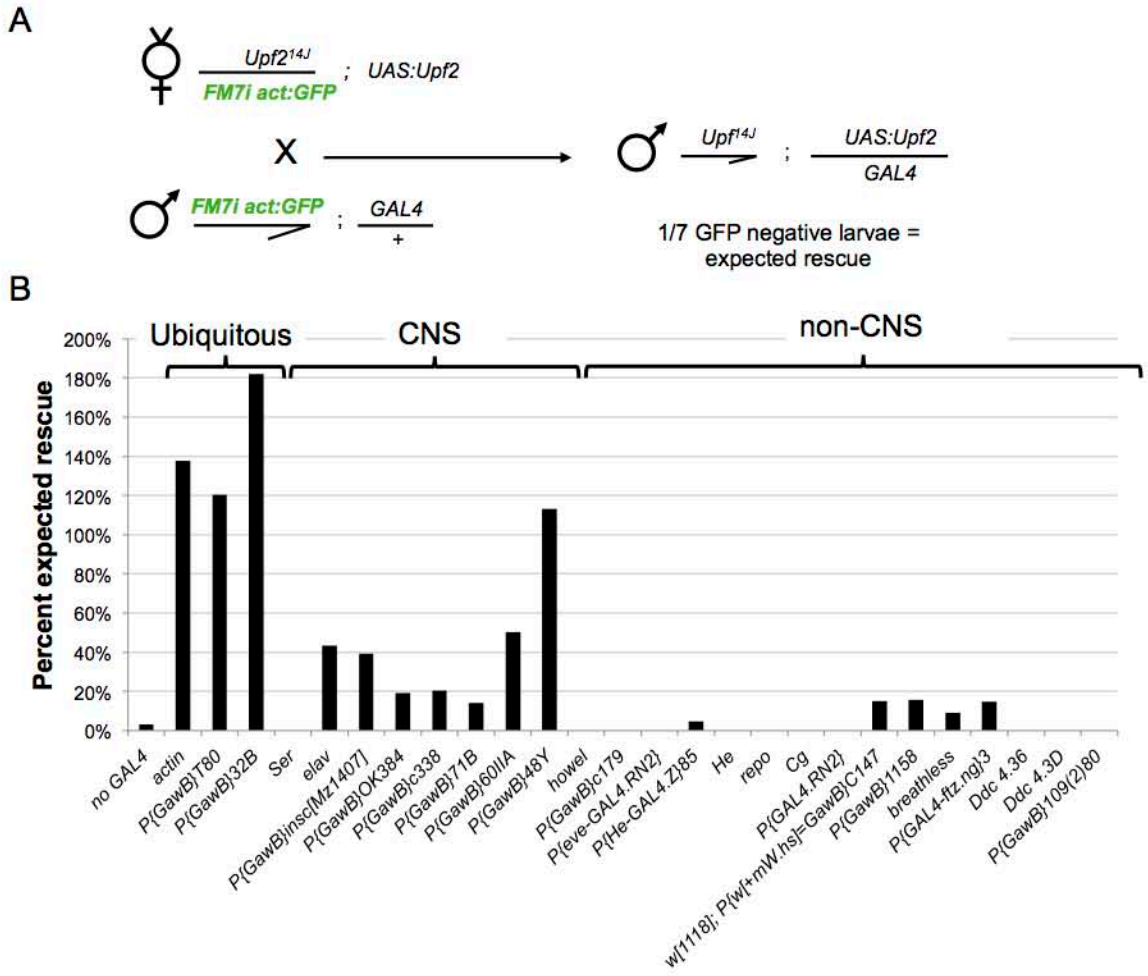


Figure 4.1: A tissue-specific rescue screen for sufficiency of NMD function: (A) Crosses used to generate genotypes for screening. (B) Proportion of expected $Upf2^{14J}/Y$ larvae assuming full rescue. Expected rescue is $1/7^{th}$ total progeny.

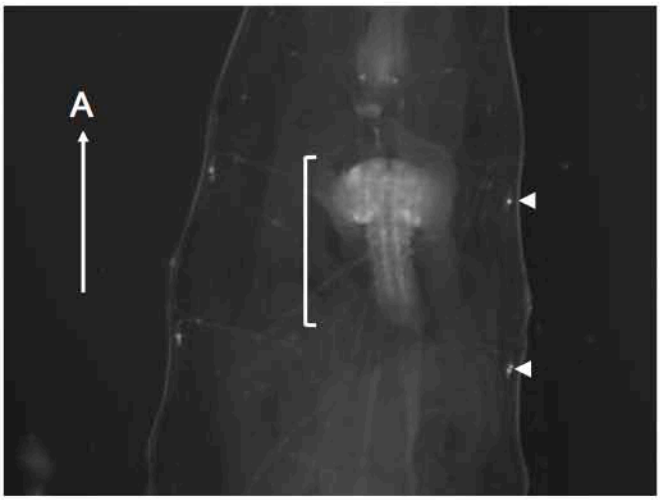


Figure 4.2: Expression pattern of the rescuing GAL4 driver *P{GawB}60IIA* as assessed by *UAS:eGFP* reporter expression. Ventral view, anterior is up. The CNS is denoted with the bracket and clusters of cell bodies belonging to the peripheral nervous system denoted with arrowheads.

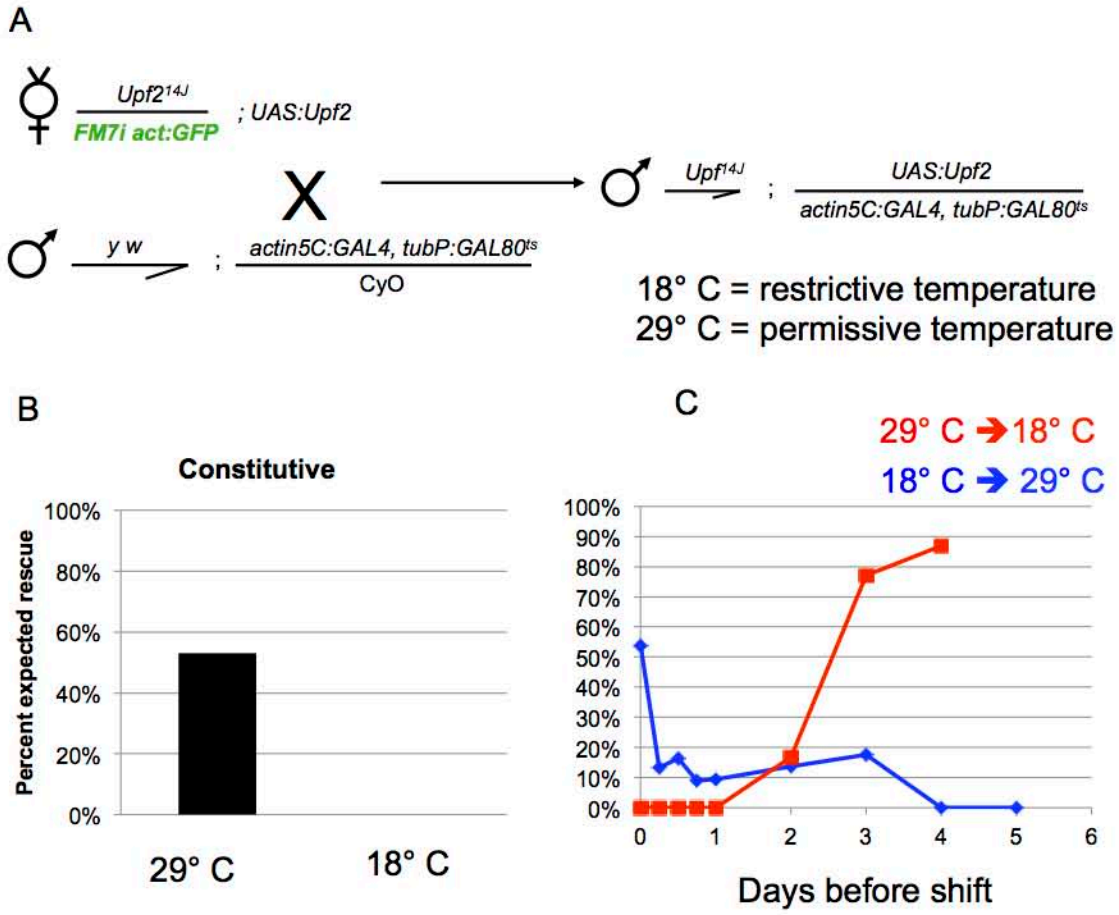


Figure 4.3: The "temperature sensitive period" of NMD function. (A) Crosses used to generate appropriate genotypes. At the permissive temperature of 29 °C, GAL80 unfolds, allowing rescuing expression from the *UAS:Upf2* cDNA. At the restrictive temperature of 18 °C, *UAS:Upf2* expression is inhibited. (B) Expected percentage of $Upf2^{14J}/Y$ adults after being held constitutively at either the permissive or restrictive temperatures. (C) Proportion of expected rescued $Upf2^{14J}/Y$ adults after either an upshift (blue) or downshift (red) at the indicated time after egg lay.

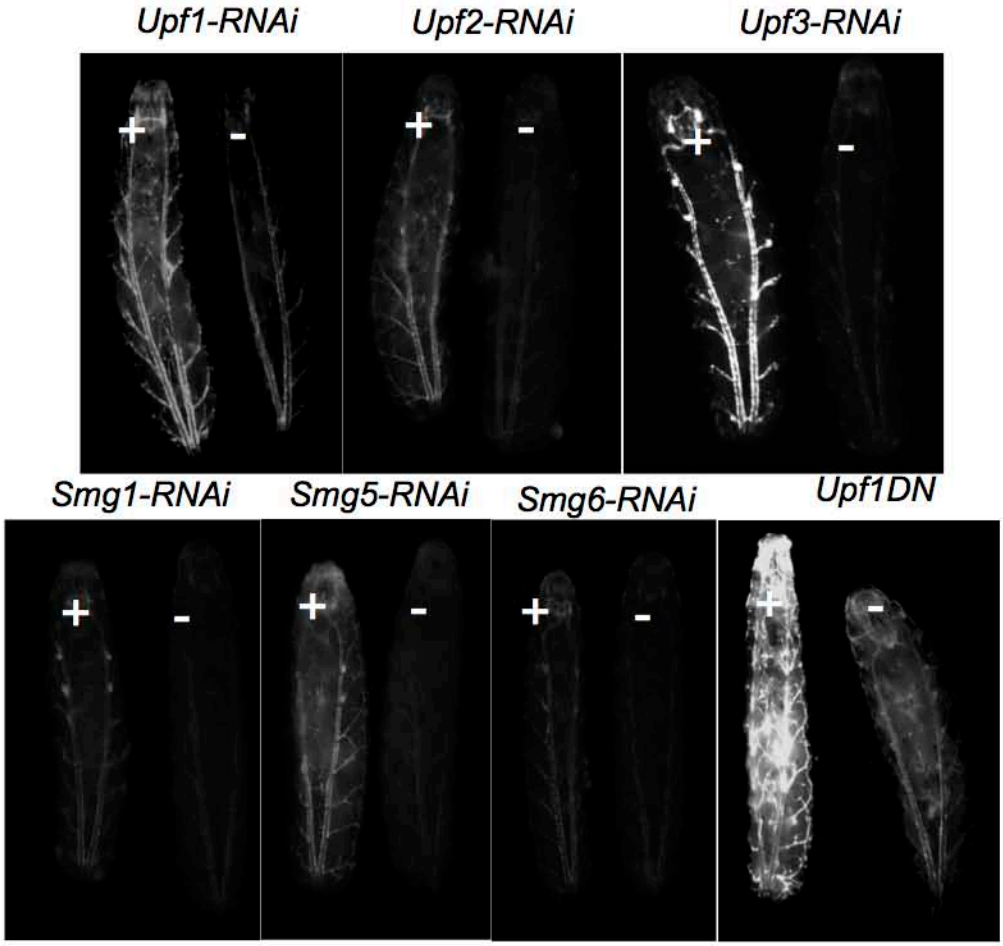


Figure 4.4 Effect of NMD-component inhibition on stabilization of a fluorescent, NMD-sensitive reporter. In this experiment, *UAS:RNAi* or *UAS:Upf1-DN* (dominant negative) transgenes are crossed to a stock with *btl:GAL4* in addition to an NMD-sensitive *UAS:eGFP SV40 3'UTR*. In all frames, the larvae on the left carries the RNAi/DN transgene indicated above (+), whereas the sibling larvae on the right carry no RNAi/DN (-).

Table 4.1: Stock numbers, genotypes, and expression patterns of GAL4 driver lines used in the tissue-specific rescue screen. GAL4 driver lines were crossed to a *UAS:eGFP, UAS:nlsDsRed* double reporter stock and F1 progeny dissected at the third larval stage. + indicates eGFP and DsRed were observed in the indicated tissue. CNS stands for central nervous system; PNS, peripheral nervous system; PV, proventriculus of the foregut; discs, imaginal discs.

		Expression										
Bloomington Stock Number	Genotype	*Epidermis*	*CNS*	*PNS*	*Trachea*	*Fat Body*	*Foregut*	*Midgut*	*Hindgut*	*Somatic Musculature*	*Dorsal Vesicle*	*Salivary Glands*
none	none	-	-	-	-	-	-	-	-	-	-	-
hsGAL4 III	w[*]; P{w[+mC]=GAL4-Hsp70.PB}89-2-1	anterior	-	N/A	-	-	+	PV and other cells	-	-	N/A	+
actin GAL4	Act5C	++	++	N/A	+	+	++	++	++	N/A	+	+
BL1878	w[*]; P{w[+mW.hs]=GawB}T80/CyO	+	++	N/A	+	++	++	++	++	N/A	N/A	+
BL1782	w[*]; P{w[+mW.hs]=GawB}32B	+	++ ocular lobes	N/A	+	+	+	++	+	N/A	N/A	-
BL6791	w[*]; P{w[+mC]=Ser-GAL4.GF}1 P{Ser-GAL4.GF}2	anterior	+	N/A	-	-	N/A	-	-	-	-	-
BL8765	P{w[+mC]=GAL4-elav.L}2/CyO	-	+	+	-	-	-	-	-	-	-	-
BL8751	w[*]; P{w[+mW.hs]=GawB}insc[Mz1407]	-	+	-	-	-	+/-	+	+	-	N/A	-
BL9448	w[*]; P{w[+mW.hs]=GawB}OK384/TM3, Sb[1]	segmental banding	+	N/A	+	+	+	Select Cells	-	N/A	-	+
BL3736	w[1118]; P{w[+mW.hs]=GawB}c338/T(2;3)ap[Xa], P{GawB}c338	+	+	-	-	-	+	+/-	-	-	N/A	+
BL1747	y[1] w[*]; P{w[+mW.hs]=GawB}71B	-	+	-	+	-	N/A	-	-	-	N/A	-
BL7029	y[1] w[*]; P{w[+mW.hs]=GawB}60IIA	very dim	+	+	-	-	N/A	select cells	-	-	N/A	-
BL4935	w[*]; P{w[+mW.hs]=GawB}48Y	-	+	-	mosaic	-	-	Select Cells	patch of cells	-	N/A	-
BL1767	w[*]; P{w[+mW.hs]=GawB}how[24B]	+	certain cells	maybe	Dorsal Trunk	+	+	+/PV	certain cells/ trachea?	N/A	N/A	+

Table 4.1 continued.

BL6450	w[*]; P{w[+mW.hs]=GawB}c179	+	ocular lobes	-	+	-	+	+	+/pig tubes	N/A	N/A	+
BL1795	?	-	subdomain of ocular lobes	N/A	-	-	-	-	-	-	N/A	+
BL8769	y[1] w[*]; P{w[+mW.hs]=GawB}109(2)80	few cells	Dim	+	Posterior trunk	weak	N/A	select cells	-	-	N/A	-
breathless GAL4	btl:GAL4	segmental banding	-	-	+++	-	-	-	-	-	-	-
BL8700	w[*]; P{w[+mC]=He-GAL4.Z}85, P{w[+mC]=UAS-GFP.nls}8	certain cells	-	N/A	-	+/-	-	arms of PV, posterior cells	-	-	-	+
BL8699	w[*]; P{w[+mC]=He-GAL4.Z}85	+	-	maybe	-	-	-	arms of PV, other cells	posterior	-	N/A	+
BL7415	w[1118]; P{w[+m*]=GAL4}repo/TM3, Sb[1]	-	-	N/A	-	-	-	broad, weak	-	-	N/A	+
BL7011	w[1118]; P{w[+mC]=Cg-GAL4.A}2	-	Very Dim	-	-	++	N/A	-	-	-	N/A	-
BL7470	y[1] w[*]; P{w[+mC]=eve-GAL4.RN2}E											-
BL6979	w[1118]; P{w[+mW.hs]=GawB}C147	certain cells	few	-	Dim	-	+	-	-	-	N/A	+
BL7028	P{w[+mW.hs]=GawB}1158x, w[*]	-	++	N/A	-	-	Dim	-	-	-	N/A	-
BL1822	y[1] w[*]; P{w[+mC]=GAL4-Hsp70.PB}31-1/T(2;3)B3, CyO: TM6B, Tb[+]	-	mosaic, but many cells	-	-	-	-	+	++	-	N/A	+
Bl8767	w[*]; P{w[+m*]=GAL4-ftz.ng}3/TM3, P{w[+mC]=act-lacZ.B}JT1, Sb[1]											-
BL7471	w[*]; P{w[+mC]=eve-GAL4.RN2}E, P{w[+mC]=UAS-tau-lacZ.B}3	-	-	-	-	-	-	-	-	-	-	-
BL7009	w[1118]; P{w[+mC]=Ddc-GAL4.L}4.36	+	few cells	-	-	-	+ certain cells	-	-	-	N/A	-
BL7010	w[1118]; P{w[+mC]=Ddc-GAL4.L}4.3D	+	certain cells in "shaft"	-	-	-	+ PV	-	-	-	-	-
BL8768	y[1] w[*]; P{w[+mW.hs]=GawB}109(2)80, P{w[+mC]=UAS-mCD8::GFP.L}LL5	some cells along body. NMJ?	few cells at base of ocular lobe	see epidermis maybe	-	+	+ neuronal	+ neuronal	+ neuronal	innervated?	N/A	

Table 4.1 continued.

BL7471	w[*]; P{w[+mC]=eve-GAL4.RN2}E, P{w[+mC]=UAS-tau-lacZ.B}3	-	-	-	-	-	-	-	-	-	-	-
BL7009	w[1118]; P{w[+mC]=Ddc-GAL4.L}4.36	+	few cells	-	-	-	+ certain cells	-	-	-	N/A	-
BL7010	w[1118]; P{w[+mC]=Ddc-GAL4.L}4.3D	+	certain cells in "shaft"	-	-	-	+ PV	-	-	-	-	-
BL8768	y[1] w[*]; P{w[+mW.hs]=GawB}109(2)80, P{w[+mC]=UAS-mCD8::GFP.L}LL5	some cells along body. NMJ?	few cells at base of ocular lobe	see epidermis maybe	-	+	+ neuronal	+ neuronal	+ neuronal	innervated?	N/A	

Table 4.2: Primers used in this study.

Primer name	Sequence
Upf1RNAi(a) Left Forward	GCGGCCGCTTCGCACATCATCAACCATC
Upf1RNAi(a) Left Reverse	CTCGAGCATCCTCGTAGCGGAGTAGC
Upf1RNAi(a) Right Forward	TCTAGATTCGCACATCATCAACCATC
Upf1RNAi(a) Right Reverse	GGTACCCATCCTCGTAGCGGAGTAGC
Upf1RNAi(b) Left Forward	GCGGCCGCTGGTCCCTACGGAAACTCA
Upf1RNAi(b) Left Reverse	CTCGAGTTCCCAGCTTGTTGGTCTTC
Upf1RNAi(b) Right Forward	TCTAGATGGTCCCTACGGAAACTCAC
Upf1RNAi(b) Right Reverse	GGTACCTTCCCAGCTTGTTGGTCTTC

References

Avery, P., Vicente-Crespo, M., Francis, D., Nashchekina, O., Alonso, C.R., and Palacios, I.M. (2011). Drosophila Upf1 and Upf2 loss of function inhibits cell growth and causes animal death in a Upf3-independent manner. Rna *17*, 624–638.

Bruno, I.G., Karam, R., Huang, L., Bhardwaj, A., Lou, C.H., Shum, E.Y., Song, H.-W., Corbett, M.A., Gifford, W.D., Gécz, J., et al. (2011). Identification of a microRNA that activates gene expression by repressing nonsense-mediated RNA decay. Mol. Cell *42*, 500–510.

Chapin, A., Hu, H., Rynearson, S.G., Hollien, J., Yandell, M., and Metzstein, M.M. (2014). In vivo determination of direct targets of the nonsense mediated decay pathway in Drosophila. G3 *4*, 485–496.

Colak, D., Ji, S.-J., Porse, B.T., and Jaffrey, S.R. (2013). Regulation of axon guidance by compartmentalized nonsense-mediated mRNA decay. Cell *153*, 1252–1265.

Edgar, B.A. (2006). How flies get their size: genetics meets physiology. Nat. Rev. Genet. *7*, 907–916.

Giorgi, C., Yeo, G.W., Stone, M.E., Katz, D.B., Burge, C., Turrigiano, G., and Moore, M.J. (2007). The EJC factor eIF4AIII modulates synaptic strength and neuronal protein expression. Cell *130*, 179–191.

Long, A.A., Mahapatra, C.T., Woodruff, E.A., Rohrbough, J., Leung, H.-T., Shino, S., An, L., Doerge, R.W., Metzstein, M.M., Pak, W.L., et al. (2010). The nonsense-mediated decay pathway maintains synapse architecture and synaptic vesicle cycle efficacy. J. Cell. Sci. *123*, 3303–3315.

McBrayer, Z., Ono, H., Shimell, M., Parvy, J.-P., Beckstead, R.B., Warren, J.T., Thummel, C.S., Dauphin-Villemant, C., Gilbert, L.I., and O'Connor, M.B. (2007). Prothoracicotropic hormone regulates developmental timing and body size in Drosophila. Dev Cell *13*, 857–871.

Metzstein, M.M., and Krasnow, M.A. (2006). Functions of the nonsense-mediated mRNA decay pathway in Drosophila development. PLoS Genet. *2*, e180.

Neu-Yilik, G., Gehring, N., Hentze, M., and Kulozik, A. (2004). Nonsense-mediated mRNA decay: from vacuum cleaner to Swiss army knife. Genome Biol. *5*, 218.

Rehwinkel, J., Letunic, I., Raes, J., Bork, P., and Izaurralde, E. (2005). Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. Rna *11*, 1530–1544.

Tarasoff, M., and Suzuki, D.T. (1970). Temperature-sensitive mutations in drosophila melanogaster. VI. Temperature effects on development of sex-linked recessive lethals. Dev. Biol. *23*, 492–509.

Tarpey, P.S., Raymond, F.L., Nguyen, L.S., Rodriguez, J., Hackett, A., Vandeleur, L., Smith, R., Shoubridge, C., Edkins, S., Stevens, C., et al. (2007). Mutations in UPF3B, a member of the nonsense-mediated mRNA decay complex, cause syndromic and nonsyndromic mental retardation. Nat. Genet. *39*, 1127–1133.

CHAPTER 5

CONCLUSIONS AND FUTURE DIRECTIONS

The nonsense mediated mRNA decay (NMD) pathway is a conserved, eukaryotic cellular surveillance pathway originally identified for its role in the rapid turnover of transcripts harboring nonsense mutations (Belasco, 1993). Subsequent investigation has revealed that NMD can also target and degrade other kinds of aberrant RNAs, in addition to wild-type, error-free messages (He et al., 1993). Additionally, function of the NMD pathway is essential for numerous biological processes across eukaryota and is required for overall viability in more complex organisms (Hwang and Maquat, 2011). Taken together, these observations evoke a model wherein NMD mediates essential functions through destabilization of native substrate RNAs. However, this model has not been completely tested and many questions surrounding the function of NMD remain unanswered. As described in this thesis, we have used the model genetic organism *Drosophila melanogaster* to address such questions as: What portion of the transcriptome is subject to NMD-mediated regulation? What *cis*-acting features sensitize transcripts to NMD-mediated decay? What kinds of RNAs are most critically targeted; nonfunctional/aberrant transcripts, or functional/error-free RNAs? Are there specific target genes that are critically regulated by NMD?

Chapter 2 describes experiments designed to identify target transcripts of NMD in intact animals. Previous research has often defined NMD target genes as those transcripts that are upregulated in backgrounds where NMD has been disabled. In *Drosophila* L3 larvae, we find that a significant portion of the transcriptome (4%) is upregulated in our representative NMD mutant, $Upf2^{25G}$. This result is in agreement with previous reports in *Drosophila* S2 cell culture, as well as in other model organisms, which indicate similar levels of upregulation when *Upf2* is inhibited (Rehwinkel et al., 2005; Weischenfeldt et al., 2012). Also like other organisms, we observe that many ontological categories are represented in our list of upregulated genes, implying that NMD regulates a range of biological processes. One overrepresented category is composed of genes involved in the defense response, suggesting that NMD functions to suppress the response to pathogens. This observation mirrors what has been described in *Arabidopsis*, where NMD mutants display an upregulation of pathogen response genes. This upregulation appears to be physiologically relevant, as *Arabidopsis* NMD mutants are resistant to pathogen infection (Riehs-Kearnan et al., 2012). Moreover, mutations that block full upregulation of defense response genes rescue other aspects of the NMD loss-of-function phenotype, including subviability.

While analysis of steady-state expression levels in NMD mutants is useful in understanding the overall role of NMD, these techniques cannot distinguish between primary and secondary effects of the mutation; i.e., between direct and indirect target mRNAs. With this consideration, we adapted a reactivation assay (Johansson et al., 2007; Maderazo et al., 2003), used previously only in yeast, to parse our list of NMD targets into direct (reactivation targets) and indirect (nonreactivation targets). Our finding that a

minority (16%) of genes upregulated in *Upf2*^{25G} appear to be direct targets mirrors observations in HeLa cells, where only 23% of genes which are upregulated following silencing of *Upf1* display a longer rate of decay compared to controls (Tani et al., 2012). This result implies that the majority of changes in gene expression following inhibition of NMD in human cells are indirect. That this study came to a similar conclusion about the portion of directly targeted transcripts found upregulated in NMD-mutant backgrounds (and by using a very different experimental approach) supports the notion that in complex organisms, NMD exerts the majority of influence over steady-state gene expression by indirect means. In yeast, a greater proportion of genes upregulated in *Upf2Δ* strains are reactivation targets (54%), suggesting that the majority of gene expression changes in this organism that occur in mutant backgrounds are direct effects of stabilization (Belasco, 1993; Johansson et al., 2007). In agreement with this conclusion is the observation that only a few genes are downregulated by inhibition of NMD in yeast, also implying that there exist far fewer indirect targets in this organism (He et al., 1993, 2003). This observation is in contrast to *Drosophila* and human cells, in which many genes are downregulated in NMD mutants (Hwang and Maquat, 2011; Rehwinkel et al., 2005; Tani et al., 2012). Thus, while dissimilar to yeast, our estimate about the proportion of primary and secondary targets is similar to findings in HeLa cells, suggesting the overall regulatory role of NMD on the *Drosophila* transcriptome may be more similar to the role of NMD in mammals.

Because the reactivation assay is not based on any other criteria, including steady-state levels in a mutant background, we also identified another class of genes in addition to direct and indirect NMD targets. This third class of genes are those identified as

reactivation targets, but are not upregulated in *Upf2*^{25G}. We refer to these "renormalized targets," as it is presumed that these genes are directly targeted, but also subject to compensatory downregulation such that the steady-state levels are not divergent from wild-type. Of our reactivation targets, only 19% are upregulated in *Upf2*^{25G}*/Y,* indicating that 81% are renormalized. This class of target gene is rarely discussed in the NMD literature, but has been observed in yeast, where 32% of reactivation targets are not upregulated in NMD mutants (Johansson et al., 2007; Rehwinkel et al., 2005; Weischenfeldt et al., 2012). Recent observations in HeLa cells indicate that only 10% of transcripts that are stabilized by silencing of *Upf1* are also upregulated (Riehs-Kearnan et al., 2012; Tani et al., 2012). This latter observation again mirrors our results in *Drosophila* and bolsters the notion that the majority of directly targeted transcripts are in fact renormalized such that no expression level change is detected at steady-state.

That we find so many renormalized target genes could be explained by a few scenarios. One possibility is that renormalized target genes are positively transcriptionally regulated by processes mediated by NMD itself or individual NMD factors. For instance, NMD genes could act as transcriptional co-activators that are necessary for the expression of renormalized target genes. Thus, in NMD mutants, while renormalized target mRNAs are more stable, their rate of transcription is also decreased, which results in minimal deviation in expression level from wild-type. While speculative, this scenario has been proposed when considering other RNA decay processes. For example, the 5' to 3' exonuclease *Xrn1* has been shown to promote the transcription of many of its decay target. In *Xrn1* mutants, while the large majority of genes are stabilized, there is little overall change in global transcript levels (Haimovich et al., 2013; Johansson

et al., 2007; Maderazo et al., 2003). Supporting the possibility that NMD components could act in a similar way, NMD factors have been shown to have nuclear functions. For instance, UPF1 has been show to shuttle between the nuclease and cytoplasm (Tani et al., 2012; Varsally and Brogna, 2012), and deletion of *Upf1* has been shown to effect transcriptional kinetics of NMD decay targets (de Turris et al., 2011).

Our overall conclusions that a minority of genes upregulated in *Upf2*^{25G}*/Y* are direct targets and that the expression levels of a minority of direct targets are changed in *Upf2*^{25G}*/Y* mutant could have broad implications in the study of the *cis*-acting sequence features that sensitize transcripts to NMD. In many studies, features such as a long 3' UTR are enriched, but imperfectly correlated with NMD target status (Behm-Ansmant et al., 2007; Eberle et al., 2009; Hansen et al., 2009; Kebaara and Atkin, 2009; Rayson et al., 2012; Singh et al., 2008; Yepiskoposyan et al., 2011). Many of these studies rely on upregulation in an NMD mutant background to define NMD targets. This definition of an NMD target, in light of our observations, which suggest that only a minority of upregulated genes are directly targeted, could explain the weak associations with putative targeting features and actual targeting status. For instance, we show that a long 3' UTR is not associated with genes upregulated in *Upf2*^{25G}*/Y*, but is associated with genes defined as reactivation targets. These data not only bolster our conclusion that the reactivation assay has identified *bona fide* direct targets of NMD, but also may help explain the weak associations with 3' UTR length in other studies that use upregulation to define NMD targets. However, we still observe that many of our reactivation targets do not have long 3' UTRs, suggesting that other sensitizing, *cis*-acting features of native NMD substrates remain undiscovered. In the future, it would be interesting to examine more properties of

our reactivation targets. For instance, our deep sequencing data could be used to examine if reactivation targets are enriched in genes that are inefficiently spliced, as such genes have been shown to be directly targeted by NMD. This hypothesis could be tested by quantifying reads mapping to pre-RNAs versus RNAs and comparing this measurement between wild-type and *Upf2*^{25G}*/Y.*

In Chapter 3 of this dissertation, we more closely examine the 24 genes that are both upregulated in *Upf2*^{25G}*/Y* and were identified as reactivation targets. We refer to this group as "critical target genes" because one or more of these genes are likely to mediate the NMD loss-of-function phenotype in *Drosophila*. To help narrow our focus to only the most important of these 24 genes, we first cross-referenced our list of critical targets with the results of a deficiency rescue screen conducted previously in the Metzstein lab. This screen was designed to identify physiologically relevant target transcripts. From this screen, a small number of genomic regions were identified that are predicted to contain important mediators of the NMD loss of function phenotype. One of these regions contains the critical target gene *Gadd45,* which we confirm is a direct NMD target in intact larvae by demonstrating that its long 3' UTR is sufficient to render a reporter RNA NMD-sensitive. We then show that disruption of the *Gadd45* CDS can partially rescue the subviability of *Upf2*^{25G}*/Y* mutants and the L2-arrest associated with *Upf2*^{14J}*/Y* mutants. These results establish that upregulation of *Gadd45* interferes with normal development and mediates at least part of the NMD mutant phenotype. These results establish *Gadd45* as a physiologically relevant critical target gene. To our knowledge, this is the first such gene identified in metazoa.

In support of our conclusion that *Gadd45* is a physiologically relevant NMD

target gene, we also show that disruption of *Gadd45's* obligate downstream signaling partner, *Mekk1,* can also rescue $Upf2^{25G}/Y$ subviability. Moreover, we also observe that six of the 13 known transcriptional targets of *Mekk1* are found as indirect targets of NMD. Taken together, these observations fit into a model wherein NMD functions to negatively regulate levels of *Gadd45*, such that in strong NMD mutants, the stabilization of *Gadd45* leads to inviability through the activation of MEKK1.

Data from other labs indicate that the negative regulation of *Gadd45* by the NMD pathway is evolutionarily conserved and that this regulation could have physiologically relevant consequences in other lineages. For instance, *Gadd45β*, a mammalian homologue of *Drosophila Gadd45*, is a direct target of NMD, possibly via its 3' UTR-localized intron (Viegas et al., 2007). Similar to *Drosophila*, this targeting could also have relevant functions in mammals as phenotypes associated with *Gadd45* activation in both these model systems mirror those of NMD loss of function. For instance *Gadd45* activation has been shown to promote cell cycle arrest and apoptosis (Jin et al., 2002; Mak and Kültz, 2004; Peretz et al., 2007; Takekawa and Saito, 1998), two process that are also stimulated by NMD inhibition in both systems (Avery et al., 2011; Azzalin et al., 2007; McIlwain et al., 2010; Medghalchi et al., 2001; Rehwinkel et al., 2006). A particularly striking example of these similarities is observed in *Drosophila*, where both *Gadd45* overexpression and NMD loss of function specifically in the female somatic germline both result in the same eggshell pattering defect, fusion of the dorsal appendages (Avery et al., 2011; Peretz et al., 2007).

While our data do support the model that *Gadd45* is a true physiologically relevant critical target of NMD, our conclusions are limited because our data, including

sequence analysis of *Gadd45*^{54C}, indicate that this allele is likely hypomorphic. In the future, alternative strategies, or a new round of TALEN mutagenesis, must be used to obtain null alleles of *Gadd45*. One approach would be to repeat the same TALEN strategy, but use a different set of HRMA primers that may be able to detect larger and/or different mutations. Three other primer pairs have been tested that generate amplicons of slightly different size around the predicted TALEN cut site, and all three amplicons appear to be amenable to the HRMA procedure. One possible reason that we did not obtain null alleles using the allele recovery strategy outlined in Chapter 3 is that *Gadd45* null mutants may be lethal. This situation would create a bias when selecting injected G_0 animals to follow in the allele recovery process as only G_0s that were both fertile and produced a shift from the wild-type melt curve in the HRMA were selected for allele recovery. For instance, if *Gadd45* disruption is in fact lethal or leads to infertility, the mosaic G_0s that harbor enough mutations in *Gadd45* to detect in HRMA could be inadvertently selected to carry mild disruptions of *Gadd45* or only harbor strong mutations in somatic cells. If the TALEN strategy is repeated, more G_0s, especially those that have no detectable shift in HRMA melt-curve, should used for propagation. Alternatively, different strategies such as imprecise P-element excision could be attempted.

In the future, it will also be necessary to characterize any suitable *Gadd45* alleles in terms of any mutant phenotype they might bestow that is independent of NMD pathway function. Possible phenotypes that could result from *Gadd45* dysfunction would likely be reminiscent of either *Mekk1* loss of function, or defects in JNK or p38 signaling, such as sensitivity to high temperature or high osmolality (Inoue et al., 2001). This could

be done by assaying viability in response to these stresses or through molecular examination of MAPK pathway readouts such as p38 phosphorylation, which is reduced compared to wild-type in *Mekk1* mutants under stress (Inoue et al., 2001). *Mekk1* mutants are also defective in the activation of apoptosis (Kang et al., 2012), a defect that could be assayed in *Gadd45* mutants using similar techniques to those used in Kang et al., (2012). *Mekk1* mutants are also defective in the transcriptional response to wounding (Brun et al., 2006), a process that activates *Gadd45* and, in a Mekk1-dependent manner, upregulates the *Turandot* genes of the humoral immune system. Therefore, it would be predicted that *Gadd45* mutants are defective in at least the transcriptional response to wounding, if not functionally compromised for this process.

It will also be important in the future to examine all aspects of the *NMD: Gadd45* double null mutant phenotype. For instance, one would predict that the small clone size phenotype of NMD mutant tissue (Avery et al., 2011; Metzstein and Krasnow, 2006) would also be rescued by *Gadd45* mutations. The partial rescue of the $Upf2^{14J}$ lethal phase by $Gadd45^{54C}$ is likely not due to the cell cycle defects observed in NMD mutants, as cell proliferation is not required for viability in *Drosophila* larvae. Therefore, *NMD; Gadd45* double mutant mitotic clones should be examined in proliferative tissue such as the imaginal discs or follicular epithelium, as has been done for NMD single mutants (Avery et al., 2011). We predict that disruption of *Gadd45* will rescue this small clone size. If such a result is obtained, it would suggest that the pro-apoptotic signaling observed in NMD mutant cells is triggered by *Gadd45* mRNA stabilization.

Another experiment to test the role of *Gadd45/Mekk1* activation in NMD mutants would be to examine genes that are upregulated by *Mekk1* in response to wounding. As

activation of *Mekk1* in NMD mutant is likely mediated through *Gadd45*, the expression levels of genes such as *TotA* (which is both an *Mekk1* and NMD-mediated transcript) are predicted to be reduced in NMD/*Gadd45* double mutants compared to NMD single mutants. Ideally, it would also be helpful to examine global mRNA levels between NMD, *Gadd45,* and *NMD/Gadd45* double mutants. This would allow for the assessment of the proportion of secondary NMD targets that are upregulated downstream of *Gadd45* stabilization in NMD mutants. As *NMD; Gadd45* double null mutants will likely be rescued for certain NMD defects, such an analysis would be of use in characterizing the cellular defects in NMD mutants. For example, the expression levels of any genes that are rescued by *Gadd45* disruption in NMD mutants could indicate what kinds of cellular pathways are critically modulated in NMD mutants.

To examine the role of MAPK activation in the NMD mutant phenotype, it will also be necessary to repeat the experiments described above using NMD:*Gadd45* double mutants with NMD and JNK pathway or NMD and p38 pathway mutants. Such analysis would be the most direct test of the model that NMD mutants succumb to over activation of MAPK signaling. These experiment would be challenging for many reasons, such as the observation that complete disruption of JNK or p38 signaling results in a pleotropic phenotype that includes organismal inviability. Thus, clonal techniques would likely be needed to generate the desired double mutant cell populations and assess any MAPK mutant modulation of the clone size defect of NMD mutants.

While consistent with our model, the interpretation of our data that $Mekk1^{Ur36}$ rescues $Upf2^{25G}$ has limitations. Primarily, no suitable wild-type control chromosome exists to compare to the two $Mekk1^{Ur36}$ stocks. Also, confirming the precise nature of the

deletion in these stocks has been challenging. gDNA from these stocks does fail to amplify a PCR band consistent with wild-type *Mekk1*, but no assay in our lab has yet demonstrated that our lab stocks harbor the deletion reported by Inoue et al. (2001). Generation of new loss-of-function alleles of *Mekk1*, with appropriate wild-type controls, would be an ideal first step toward studying the relationship between NMD and *Mekk1*.

In Chapter 4 of this dissertation, we present data suggesting that function of the NMD pathway in neuronal tissue is sufficient to rescue the L2/L3 molting defect in *Upf2* null mutants. It is currently unclear how this result might relate to our finding that targeting of *Gadd45* is necessary for viability beyond the second larval instar. A simple model would be that in NMD mutants, *Gadd45* upregulation specifically in neuronal tissue is lethal. However, expression of a *UAS:Gadd45* cDNA using the same GAL4 drivers that rescue $Upf2^{14J}$ does not result in any lethality (data not shown), whereas global overexpression of the same cDNA is lethal (Peretz et al., 2007). These results suggest that upregulation of *Gadd45* in the CNS or neuronal tissue is not the cause of inviability in NMD mutants. One speculative possibility is related to the model described in Chapter 4 that states that the rescue of $Upf2^{14J}$ in the CNS is due to a gain-of-function phenotype that does not require that NMD be essential in the CNS. In this model, NMD defects in nonneuronal tissue disrupt the L2/L3 molt, and activation of NMD in the CNS drives a second pathway (possibly related to release of PTTH from neuroendocrine cells) that drives molting. Our observations that *Gadd45* overexpression in the CNS is not incompatible with viability fit into this model, as *Gadd45* need not be critically targeted in the CNS. Instead, we predict that NMD mutants succumb to *Gadd45* overexpression in multiple, nonneuronal cell types, and expression of NMD in the CNS activates a non-

Gadd45 mediated pro-molting pathway. The model could be supported by experiments that demonstrate that NMD function is not required in neuronal tissue.

Lastly, it is likely that *Gadd45/Mekk1* does not mediate the entire phenotype of NMD loss-of-function mutants. In the future, continued analysis of our 24 critical target genes could yield further insights into the physiological and developmental roles of NMD function. For instance, a systematic approach could be used to analyze any functional consequences to the targeting of these genes by the NMD pathway. RNAi could be used to ask if individually inhibiting the function of the 23 remaining genes in a NMD-mutant background can rescue some aspect of NMD function. Additionally, part of the logic for using a hypomorphic allele of *Upf2* in the deficiency rescue screen was to identify regions that can rescue different aspects of the NMD mutant phenotype. Most of our transcriptome analysis was focused on late larval stages, but it is known that NMD acts earlier in development and is predicted to function later in development as well. Thus, continued analysis of the genes identified by this screen could aid in our understanding of NMD at all times in development and in all environmental conditions. For instance, the deficiency rescue screen identified a region around the gene *Arc1* that we believe could be an important indirect target of NMD function. *Xrp1*, a critical target gene, is also uncovered by the same deficiencies that uncover *Mekk1* and rescue $Upf2^{25G}$. While *Mekk1* likely plays an important role in NMD-mediated gene regulation, this does not preclude *Xrp1* also having important NMD-related functions. Activation of *Xrp1* leads to growth arrest (Akdemir et al., 2007), as does NMD loss of function, suggesting these genes could also function downstream of NMD, possibly along with *Gadd45*, to initiate growth arrest and apoptosis when NMD is inhibited.

In summary, we have presented data that suggest that NMD has important functions over the course of *Drosophila* development. Our data support a model similar to the Swiss Army Knife model of critical NMD function, where NMD critically regulates a small number of direct target genes, one of which is *Gadd45*. Direct negative regulation of *Gadd45* is critical for suppression of *Mekk1*/MAPK signaling and development past the L2/L3 larval molt. Continued analysis of this model, as well as of the 23 remaining critical target genes, will yield further insights into NMD function.

References

Akdemir, F., Christich, A., Sogame, N., Chapo, J., and Abrams, J.M. (2007). p53 directs focused genomic responses in Drosophila. Oncogene *26*, 5184–5193.

Avery, P., Vicente-Crespo, M., Francis, D., Nashchekina, O., Alonso, C.R., and Palacios, I.M. (2011). Drosophila Upf1 and Upf2 loss of function inhibits cell growth and causes animal death in a Upf3-independent manner. Rna *17*, 624–638.

Azzalin, C.M., Reichenbach, P., Khoriauli, L., Giulotto, E., and Lingner, J. (2007). Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. Science *318*, 798–801.

Behm-Ansmant, I., Kashima, I., Rehwinkel, J., Saulière, J., Wittkopp, N., and Izaurralde, E. (2007). mRNA quality control: an ancient machinery recognizes and degrades mRNAs with nonsense codons. FEBS Lett. *581*, 2845–2853.

Belasco, J.G. (1993). Control of messenger RNA stability (Waltham: Academic Press).

Brun, S., Vidal, S., Spellman, P., Takahashi, K., Tricoire, H., and Lemaitre, B. (2006). The MAPKKK Mekk1 regulates the expression of Turandot stress genes in response to septic injury in Drosophila. Genes Cells *11*, 397–407.

de Turris, V., Nicholson, P., Orozco, R.Z., Singer, R.H., and Muhlemann, O. (2011). Cotranscriptional effect of a premature termination codon revealed by live-cell imaging. Rna *17*, 1–14.

Eberle, A.B., Lykke-Andersen, S., Mühlemann, O., and Jensen, T.H. (2009). SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. Nat. Struct. Mol. Biol. *16*, 49–55.

Haimovich, G., Medina, D.A., Causse, S.Z., Garber, M., Millán-Zambrano, G., Barkai,

O., Chávez, S., Pérez-Ortín, J.E., Darzacq, X., and Choder, M. (2013). Gene expression Is circular: factors for mRNA degradation also foster mRNA synthesis. Cell *153*, 1000–1011.

Hansen, K.D., Lareau, L.F., Blanchette, M., Green, R.E., Meng, Q., Rehwinkel, J., Gallusser, F.L., Izaurralde, E., Rio, D.C., Dudoit, S., et al. (2009). Genome-wide identification of alternative splice forms down-regulated by nonsense-mediated mRNA decay in Drosophila. PLoS Genet. *5*, e1000525.

He, F., Peltz, S.W., Donahue, J.L., Rosbash, M., and Jacobson, A. (1993). Stabilization and ribosome association of unspliced pre-mRNAs in a yeast upf1- mutant. Proc. Natl. Acad. Sci. USA *90*, 7034–7038.

He, F., Li, X., Spatrick, P., Casillo, R., Dong, S., and Jacobson, A. (2003). Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5“ to 3” mRNA decay pathways in yeast. Mol. Cell *12*, 1439–1452.

Hwang, J., and Maquat, L.E. (2011). Nonsense-mediated mRNA decay (NMD) in animal embryogenesis: to die or not to die, that is the question. Curr. Opin. Genet. Dev. *21*, 422–430.

Inoue, H., Tateno, M., Fujimura-Kamada, K., Takaesu, G., Adachi-Yamada, T., Ninomiya-Tsuji, J., Irie, K., Nishida, Y., and Matsumoto, K. (2001). A Drosophila MAPKKK, D-MEKK1, mediates stress responses through activation of p38 MAPK. Embo J. *20*, 5421–5430.

Jin, S., Tong, T., Fan, W., Fan, F., Antinore, M.J., Zhu, X., Mazzacurati, L., Li, X., Petrik, K.L., Rajasekaran, B., et al. (2002). GADD45-induced cell cycle G2-M arrest associates with altered subcellular distribution of cyclin B1 and is independent of p38 kinase activity. Oncogene *21*, 8696–8704.

Johansson, M.J.O., He, F., Spatrick, P., Li, C., and Jacobson, A. (2007). Association of yeast Upf1p with direct substrates of the NMD pathway. Proc. Natl. Acad. Sci. USA *104*, 20872–20877.

Kang, M.-J., Chung, J., and Ryoo, H.D. (2012). CDK5 and MEKK1 mediate pro-apoptotic signalling following endoplasmic reticulum stress in an autosomal dominant retinitis pigmentosa model. Nat. Cell Biol. *14*, 409–415.

Kebaara, B.W., and Atkin, A.L. (2009). Long 3'-UTRs target wild-type mRNAs for nonsense-mediated mRNA decay in Saccharomyces cerevisiae. Nucleic Acids Res. *37*, 2771–2778.

Maderazo, A.B., Belk, J.P., He, F., and Jacobson, A. (2003). Nonsense-containing mRNAs that accumulate in the absence of a functional nonsense-mediated mRNA decay pathway are destabilized rapidly upon its restitution. Mol. Cell Biol. *23*, 842–851.

Mak, S.K., and Kültz, D. (2004). Gadd45 proteins induce G2/M arrest and modulate

apoptosis in kidney cells exposed to hyperosmotic stress. J. Biol Chem. *279*, 39075–39084.

McIlwain, D.R., Pan, Q., Reilly, P.T., Elia, A.J., McCracken, S., Wakeham, A.C., Itie-Youten, A., Blencowe, B.J., and Mak, T.W. (2010). Smg1 is required for embryogenesis and regulates diverse genes via alternative splicing coupled to nonsense-mediated mRNA decay. Proc. Natl. Acad. Sci. USA *107*, 12186–12191.

Medghalchi, S.M., Frischmeyer, P.A., Mendell, J.T., Kelly, A.G., Lawler, A.M., and Dietz, H.C. (2001). Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. Hum. Mol. Genet. *10*, 99–105.

Metzstein, M.M., and Krasnow, M.A. (2006). Functions of the nonsense-mediated mRNA decay pathway in Drosophila development. PLoS Genet. *2*, e180.

Peretz, G., Bakhrat, A., and Abdu, U. (2007). Expression of the Drosophila melanogaster GADD45 homolog (CG11086) affects egg asymmetric development that is mediated by the c-Jun N-terminal kinase pathway. Genetics *177*, 1691–1702.

Rayson, S., Arciga-Reyes, L., Wootton, L., De Torres Zabala, M., Truman, W., Graham, N., Grant, M., and Davies, B. (2012). A role for nonsense-mediated mRNA decay in plants: pathogen responses are induced in Arabidopsis thaliana NMD mutants. PLoS ONE *7*, e31917.

Rehwinkel, J., Letunic, I., Raes, J., Bork, P., and Izaurralde, E. (2005). Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. Rna *11*, 1530–1544.

Rehwinkel, J., Natalin, P., Stark, A., Brennecke, J., Cohen, S.M., and Izaurralde, E. (2006). Genome-wide analysis of mRNAs regulated by Drosha and Argonaute proteins in Drosophila melanogaster. Mol. Cell Biol. *26*, 2965–2975.

Riehs-Kearnan, N., Gloggnitzer, J., Dekrout, B., Jonak, C., and Riha, K. (2012). Aberrant growth and lethality of Arabidopsis deficient in nonsense-mediated RNA decay factors is caused by autoimmune-like response. Nucleic Acids Res. *40*, 5615–5624.

Singh, G., Rebbapragada, I., and Lykke-Andersen, J. (2008). A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay. Plos Biol. *6*, e111.

Takekawa, M., and Saito, H. (1998). A family of stress-inducible GADD45-like proteins mediate activation of the stress-responsive MTK1/MEKK4 MAPKKK. Cell *95*, 521–530.

Tani, H., Imamachi, N., Salam, K.A., Mizutani, R., Ijiri, K., Irie, T., Yada, T., Suzuki, Y., and Akimitsu, N. (2012). Identification of hundreds of novel UPF1 target transcripts by direct determination of whole transcriptome stability. RNA Biol. *9*, 1370–1379.

Varsally, W., and Brogna, S. (2012). UPF1 involvement in nuclear functions. Biochem.

Soc. Trans. *40*, 778–783.

Viegas, M.H., Gehring, N.H., Breit, S., Hentze, M.W., and Kulozik, A.E. (2007). The abundance of RNPS1, a protein component of the exon junction complex, can determine the variability in efficiency of the Nonsense Mediated Decay pathway. Nucleic Acids Res. *35*, 4542–4551.

Weischenfeldt, J., Waage, J., Tian, G., Zhao, J., Damgaard, I., Jakobsen, J.S., Kristiansen, K., Krogh, A., Wang, J., and Porse, B.T. (2012). Mammalian tissues defective in nonsense-mediated mRNA decay display highly aberrant splicing patterns. Genome Biol. *13*, R35.

Yepiskoposyan, H., Aeschimann, F., Nilsson, D., Okoniewski, M., and Mühlemann, O. (2011). Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. Rna *17*, 2108–2118.